

#3
G-2700

THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of : Hiroshi TOMONAGA et al.

Filed : Concurrently herewith

For : PACKET SWITCH

Serial No. : Concurrently herewith

August 22, 2000

Assistant Commissioner of Patents
Washington, D.C. 20231

JCS46 U.S. PTO
09/643556
08/22/00

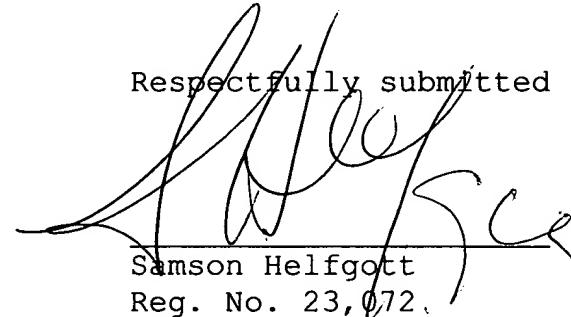


SUBMISSION OF PRIORITY DOCUMENT

S I R:

Attached herewith is Japanese patent application No. 11-235596 of August 23, 1999 whose priority has been claimed in the present application.

Respectfully submitted



Samson Helfgott
Reg. No. 23,072.

HELGOTT & KARAS, P.C.
60th FLOOR
EMPIRE STATE BUILDING
NEW YORK, NY 10118
DOCKET NO.: FUJY17.696
LHH:priority

Filed Via Express Mail
Rec. No.: EL522338187US
On: August 22, 2000
By: Lydia Gonzalez

Any fee due as a result of this paper, not covered by an enclosed check may be charged on Deposit Acct. No. 08-1634.

OP1020

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
in this Office.

出願年月日
Date of Application:

1999年 8月23日

願番号
Application Number:

平成11年特許願第235596号

願人
Applicant(s):

富士通株式会社

JC846 U.S. PRO
09/643566
08/22/00


CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 6月16日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特2000-3045888

【書類名】 特許願
【整理番号】 9900785
【提出日】 平成11年 8月23日
【あて先】 特許庁長官殿
【国際特許分類】 H04L 12/48
【発明の名称】 パケットスイッチ装置
【請求項の数】 22
【発明者】
【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株
式会社内
【氏名】 朝永 博
【発明者】
【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株
式会社内
【氏名】 松岡 直樹
【発明者】
【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株
式会社内
【氏名】 瓦井 健一
【発明者】
【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号富士通株
式会社内
【氏名】 加藤 次雄
【特許出願人】
【識別番号】 000005223
【氏名又は名称】 富士通株式会社
【代理人】
【識別番号】 100089244
【弁理士】

【氏名又は名称】 遠山 勉

【選任した代理人】

【識別番号】 100090516

【弁理士】

【氏名又は名称】 松倉 秀実

【連絡先】 03-3669-6571

【手数料の表示】

【予納台帳番号】 012092

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705606

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケットスイッチ装置

【特許請求の範囲】

【請求項1】 出力回線対応の論理的なキューを有する入力バッファメモリ手段と；

スケジューリング開始入力回線を示す第1のポインタの制御手段と；

スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；

所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；

複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；

複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；

前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；

前記パケットバッファメモリ手段のアドレスを複数パケット分の固定長ブロックに分け、アドレス管理をブロック単位に行うアドレス管理手段と；

を備えることを特徴とするパケットスイッチ装置。

【請求項2】 出力回線対応の論理的なキューを有する入力バッファメモリ手段と；

スケジューリング開始入力回線を示す第1のポインタの制御手段と；

スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；

所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；

複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；

複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモ

り手段と；

前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；

複数の出力回線にマルチキャストするために、一旦マルチキャスト用のメモリに前記パケットを書き込み、それをマルチキャスト数読み出して、それぞれを所望のFIFOメモリに振り分けを行うとき、振り分けは実際の前記パケットでなくアドレスのみで行うことを論理的に実現するために、振り分け後のアドレス毎にマルチキャスト数とそのアドレスとを保持し、前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段と；

を備えることを特徴とするパケットスイッチ装置。

【請求項3】 前記アドレス管理手段は、更にブロック内でどのシーケンス番号のどのアドレスがマルチキャストを完了したかを示すフラグにより管理することを特徴とする請求項1記載のパケットスイッチ装置。

【請求項4】 前記パケットバッファメモリ手段は、バーストアクセス時のみ高速アクセス可能な大容量メモリを前段に、高速ランダムアクセス可能なメモリを後段に配し、通常は後段のメモリのみ使用し、後段メモリが一杯になったときは前段メモリに一時的に保持し、後段メモリに空きができた時点で移し替えることを特徴とする請求項1または2記載のパケットスイッチ装置。

【請求項5】 前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、アクセスが低速な大容量メモリをパラレルに後段に配し、後段メモリへの書き込み動作は、前段メモリからパラレルに行い、後段メモリの読み出し動作は競合が起こらないキューのみを選択して読み出しを行うことを特徴とする請求項1または2記載のパケットスイッチ装置。

【請求項6】 前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、バーストアクセス時のみ高速アクセス可能な大容量メモリを用いた待ち合わせキューを後段に配し、後段メモリへの書き込み動作は、前段メモリに複数パケット溜まった時点でまとめて行い、後段メモリの読み出し動作は、複数パケットまとめて行うことの特徴とす

る請求項1または2記載のパケットスイッチ装置。

【請求項7】 前記パケットバッファメモリ手段は、複数の入力回線の固定長パケットを1つの入力回線上に確定的に時分割多重し、多重前の入力回線毎にメモリをパラレルに配置し、書き込み及び読み出しを各メモリでパラレルに行うことを特徴とする請求項1または2記載のパケットスイッチ装置。

【請求項8】 前記共通スイッチ手段は、ピットスライス構成であり、スライスされた少なくとも1つのスイッチ手段を冗長に持つことにより、スライスされたスイッチ手段毎の保守・故障時の切り替えを可能とすることを特徴とする請求項1または2記載のパケットスイッチ装置。

【請求項9】 前記パケットバッファメモリ手段は、前記共通スイッチ手段の入力側及び出力側においてそれぞれ二重化され、出力側配置の前記パケットバッファメモリ手段へのデータ分配は予め前記共通スイッチ手段に設定した方路に基づいて行うことにより、保守・故障時の切り替えを可能とすることを特徴とする請求項1または2記載のパケットスイッチ装置。

【請求項10】 前記第1及び第2のポインタの制御手段と前記要求管理制御手段と前記スケジューリング処理手段とを含むスケジューラを分散配置し、かつ前記入力バッファメモリ手段を含む入力バッファ部の隣接する前記入力バッファ部間のスケジューリング情報を選択するためのスイッチ手段を更に設け、保守・故障時の切り替えを可能とすることを特徴とする請求項1または2記載のパケットスイッチ装置。

【請求項11】 出力回線対応の論理的なキューを有する入力バッファメモリ手段と；

スケジューリング開始入力回線を示す第1のポインタの制御手段と；

スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；

所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；

複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；

複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；

前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；

前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；

前記パケットバッファメモリ手段は、バーストアクセス時ののみ高速アクセス可能な大容量メモリを前段に、高速ランダムアクセス可能なメモリを後段に配し、通常は後段のメモリのみ使用し、後段メモリが一杯になったときは前段メモリに一時的に保持し、後段メモリに空きができた時点で移し替えることを特徴とするパケットスイッチ装置。

【請求項12】 出力回線対応の論理的なキューを有する入力バッファメモリ手段と；

スケジューリング開始入力回線を示す第1のポインタの制御手段と；

スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；

所望の出力回線に対する送出要求情報を保持する要求管理体制手段と；

複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；

複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；

前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；

前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；

前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、アクセスが低速な大容量メモリをパラレルに後段に配し、後段メモリへの書き込み動作は、前段メモリからパラレルに行

い、後段メモリの読み出し動作は競合が起こらないキューのみを選択して読み出しを行うことを特徴とするパケットスイッチ装置。

【請求項13】 出力回線対応の論理的なキューを有する入力バッファメモリ手段と；

スケジューリング開始入力回線を示す第1のポインタの制御手段と；

スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；

所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；

複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；

複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；

前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；

前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；

前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、バーストアクセス時の高速アクセス可能な大容量メモリを用いた待ち合わせキューを後段に配し、後段メモリへの書き込み動作は、前段メモリに複数パケット溜まった時点でまとめて行い、後段メモリの読み出し動作は、複数パケットまとめて行うことを特徴とするパケットスイッチ装置。

【請求項14】 出力回線対応の論理的なキューを有する入力バッファメモリ手段と；

スケジューリング開始入力回線を示す第1のポインタの制御手段と；

スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；

所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；

複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；

複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；

前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；

前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；

前記パケットバッファメモリ手段は、複数の入力回線の固定長パケットを1つの入力回線上に確定的に時分割多重し、多重前の入力回線毎にメモリをパラレルに配置し、書き込み及び読み出しを各メモリでパラレルに行うことの特徴とするパケットスイッチ装置。

【請求項15】 前記共通スイッチ手段は、ビットスライス構成であり、スライスされた少なくとも1つのスイッチ手段を冗長に持つことにより、スライスされたスイッチ手段毎の保守・故障時の切り替えを可能とすることの特徴とする請求項11，12，13または14記載のパケットスイッチ装置。

【請求項16】 前記パケットバッファメモリ手段は、前記共通スイッチ手段の入力側及び出力側においてそれぞれ二重化され、出力側配置の前記パケットバッファメモリ手段へのデータ分配は予め前記共通スイッチ手段に設定した方路に基づいて行うことにより、保守・故障時の切り替えを可能とすることの特徴とする請求項11，12，13，14または15記載のパケットスイッチ装置。

【請求項17】 前記第1及び第2のポインタの制御手段と前記要求管理制御手段と前記スケジューリング処理手段とを含むスケジューラを分散配置し、かつ前記入力バッファメモリ手段を含む入力バッファ部の隣接する前記入力バッファ部間のスケジューリング情報を選択するためのスイッチ手段を更に設け、保守・故障時の切り替えを可能とすることの特徴とする請求項11，12，13，14，15または16記載のパケットスイッチ装置。

【請求項18】 ある一定速度でスケジューリング処理を行うスケジューリン

グ処理手段と；

入力回線速度分のスケジューリング速度で求められるパケットスロット時間を計測する第1のタイマ処理手段と；

出力回線速度分のスケジューリング速度で求められるパケットスロット時間を計測する第2のタイマ処理部とを備え；

ある入力回線に対するスケジューリング処理が行われたとき、前記第1のタイマ処理手段が計測を開始し、以降前記第1のタイマ処理手段が所定時間を計測する間は、その入力回線に対するスケジューリング処理を停止し、入力回線速度に応じたスケジューリングを実現することを特徴とするパケットスイッチ装置。

【請求項19】 ある出力回線に対する確定が行われたとき、その出力回線に対する前記第2のタイマ処理手段の計測を開始し、以降前記第2のタイマ処理手段が所定時間を計測する間は、同一出力回線に対する確定を停止し、同一出力回線へのトラヒック流を出力回線速度以下に抑えることを特徴とする請求項18記載のパケットスイッチ装置。

【請求項20】 前記第1及び第2のタイマ処理手段の計測開始契機をスケジューリング時あるいは確定時とせずに、固定的な時間間隔とすることを特徴とする請求項18記載のパケットスイッチ装置。

【請求項21】 パイプライン的にスケジューリング処理を行う際、スケジューリング速度と入力回線速度との比に基づく前後N段のパイプライン処理の該当入力回線に対するスケジューリングを停止し、入力回線速度に応じたスケジューリングを実現することを特徴とする請求項18記載のパケットスイッチ装置。

【請求項22】 前記第2のタイマ処理手段は、パイプライン的にスケジューリング処理を行う際、パイプライン番号と出力回線番号とに基づきパイプライン処理毎に独立に制御し、各パイプライン処理内で所定の間隔を守ることで、平均的に各出力回線速度に応じたスケジューリングを実現することを特徴とする請求項18記載のパケットスイッチ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は広帯域交換機、クロスコネクトスイッチ装置、及びルータ装置などに適用され、固定長パケット（セル）を伝送（特に、限定しないときは、伝達及び転送を含む）するパケットスイッチ装置に関し、特に大規模パケットスイッチ装置の一構成法である入力バッファ型や入出力バッファ型スイッチ装置に関する。

【0002】

【従来の技術】

近年、インターネットの爆発的な普及や大容量・高品質な情報を扱うメディアの登場により、大容量データを柔軟に扱うことのできる大規模な通信ネットワークインフラの整備に期待が寄せられている。そして、実現の鍵となる数百ギガ～数テラオーダのスイッチィング容量を持つ広帯域交換機に関心が集まっている。

【0003】

この交換機などに適用され、固定長パケットを伝送するパケットスイッチ装置において、入力方路（入力回線：HW）毎にFIFO型の单一の入力バッファメモリを有する基本的な入力バッファ型スイッチ（図1参照）は、HOL(Head Of Line)ブロッキングの問題があり、スループットが58.6%までしか上がらないことが知られている。このHOLブロッキングを回避する手法として、入力バッファ部の各入力バッファメモリを論理的に出力方路（出力回線）対応のキューに分割し、所定のアルゴリズムに従って送出権をスケジューリングすることが知られている（図2参照）。

【0004】

これにより、入力バッファ部の動作速度を低減し、汎用メモリを使用した大容量入力バッファ部の構築が可能となる。また、共通スイッチ部となるコアスイッチ（Core Switch）はバッファレスであり、ハードウェアを極小化でき、またスケジューラ（Scheduler）は分散パイプライン構成により動作速度を低減できる（図3参照）。

【0005】

ここで、入力バッファメモリは論理出方路キュー間で共用して使用されるため、キュー毎の使用アドレスを管理する必要があるが、入力バッファメモリが大容量になると、アドレス管理メモリ（FIFO）も大きくなる（図4参照）。従来

、このアドレス管理メモリの容量削減のために、ポインタを用いたアドレスリンク方式が知られている（図5参照）。

【0006】

入力バッファ部を構成するための汎用メモリとしては、比較的小容量だが高速でランダムアクセス可能なSRAMと、大容量だが高速でアクセスするためにはバーストアクセスが必要なDRAMとがある。パケット（セル）バッファメモリをDRAMにより大容量化する場合、パケット長の範囲でバーストアクセスを可能にしている（図6参照）。

【0007】

また、冗長構成としては、共通装置には二重化構成を採り、ブロック毎に交絡を設け、セレクタを通して切り替えを行うことにより、装置の信頼性を上げている（図7参照）。入力バッファ部においては、全てのパッケージカードで同一速度により読み出しを行っている。

【0008】

【発明が解決しようとする課題】

さらに、サービスの多様化に伴って、低速から高速の様々な速度のインターフェースが提供されている。このような回線速度の異なるインターフェースを共通のルータ装置やクロスコネクトスイッチ装置に効率よく混在収容することが望まれている。

【0009】

従来の混在収容の手法では、複数の低速回線を多重化してルータ装置やクロスコネクトスイッチ装置のポート速度まで速度を上げて収容するものや、回線インターフェースに備えられるバッファメモリで速度変換を行って混在収容を行っていた。しかし、前者は、比較的少ない回線インターフェースを収容する場合でも、一旦高速多重化するためのパケット多重部（MUX）が必要となり、また後者は、比較的安価に提供できる低速インターフェースに高速なバッファメモリが必要であった。

【0010】

メモリのアドレス管理については、アドレスリンク手法を用いる場合でも、バ

バッファメモリ容量分のアドレス管理メモリが必要になる。これに対し、バッファメモリ領域の一部をアドレス管理メモリとして使用する手法があるが、この手法では、パケットバッファメモリのアクセス数が増えてしまい、高速アクセスのボトルネックになる（図8参照）。

【0011】

また、マルチキャスト機能を実現する場合、複数の論理出方路キューで同一のパケットバッファメモリを使用することになるため、アドレスリンクが形成できず、アドレスリンク手法が適用できなかった（図9参照）。

【0012】

パケットバッファメモリについては、メモリを可能な限り高速に使おうとすると、パラレル化する必要がある。しかし、パケットを処理するためのパラレル度はパケット長で制限される。例えば、ATM（非同期転送モード）では53byteである。その時、全ビットパラレルで処理すると、一回のアクセスで全データを読むことになる。このように、パラレル度を増やした場合、DRAMでのバーストアクセスは適用できなくなり、高速アクセスを不可能にする（図10参照）。

【0013】

入力バッファ部の冗長構成については、二重化構成を探るため、ハードウェア量及びブロック間接続数が共に2倍となり、コストアップになる。ここで、コストダウンのために、N+1冗長構成の採用が考えられるが、入力バッファ部の後段配置の共通スイッチ部（コアスイッチ）全体で1つの固まりであり、この冗長構成の適用は難しかった。

【0014】

また、スケジューラを各入力バッファ部に分散配置する場合、スケジューリング情報の伝達路がリング状に接続されるため、入力バッファ部間のパッケージカードの一枚を抜いてしまうと、伝達情報が途切れてしまうことを免れない（図11参照）。

【0015】

本発明の第1の課題は、メモリアドレスをブロック単位で管理し、ブロック内

の個別のアドレスは書き込み時または読み出し時にキー毎に与えることでメモリ量を削減することが可能なパケットスイッチ装置を提供することにある。

【0016】

本発明の第2の課題は、少数の回線収容時でも必要となっていたパケット多重部（MUX）や低速インターフェースにおける高速バッファメモリの配置を不要とし、低速から高速まで全ての回線インターフェースを混在収容することを可能にするパケットスイッチ装置を提供することにある。

【0017】

本発明の第3の課題は、高速アクセスを実現しつつ大容量のパケット（セル）バッファメモリを構成することができるパケットスイッチ装置を提供することにある。

【0018】

本発明の第4の課題は、出力側のパケット多重分離部（DMUX）における所要バッファメモリ量を削減することを可能にするパケットスイッチ装置を提供することにある。

【0019】

【課題を解決するための手段】

上記課題を解決するために、本発明のパケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と；スケジューリング開始入力回線を示す第1のポインタの制御手段と；スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；所望の出力回線に対する送出要求情報を保持する要求管理体制と；複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；前記パケットバッファメモリ手段のアドレスを複数パケット分の固定長ブロックに分け、アドレス管理をブロック単位に行うアドレス管理手段とを備える。

【0020】

この構成において、前記アドレス管理手段は、更にブロック内でどのシーケンス番号のどのアドレスがマルチキャストを完了したかを示すフラグにより管理することもできる。

【0021】

また、本発明のパケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と；スケジューリング開始入力回線を示す第1のポインタの制御手段と；スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；複数の出力回線にマルチキャストするために、一旦マルチキャスト用のメモリに前記パケットを書き込み、それをマルチキャスト数読み出して、それぞれを所望の FIFOメモリに振り分けを行うとき、振り分けは実際の前記パケットでなくアドレスのみで行うことを論理的に実現するために、振り分け後のアドレス毎にマルチキャスト数とそのアドレスとを保持し、前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備える。

【0022】

上記各パケットスイッチ装置において、前記パケットバッファメモリ手段は、バーストアクセス時のみ高速アクセス可能な大容量メモリを前段に、高速ランダムアクセス可能なメモリを後段に配し、通常は後段のメモリのみ使用し、後段メモリが一杯になったときは前段メモリに一時的に保持し、後段メモリに空きができる時点で移し替える。

【0023】

また、前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、アクセスが低速な大容量メモリを

パラレルに後段に配し、後段メモリへの書き込み動作は、前段メモリからパラレルに行い、後段メモリの読み出し動作は競合が起こらないキューのみを選択して読み出しを行う。

【0024】

また、前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、バーストアクセス時のみ高速アクセス可能な大容量メモリを用いた待ち合わせキューを後段に配し、後段メモリへの書き込み動作は、前段メモリに複数パケット溜まった時点でまとめて行い、後段メモリの読み出し動作は、複数パケットまとめて行う。

【0025】

さらに、前記パケットバッファメモリ手段は、複数の入力回線の固定長パケットを1つの入力回線上に確定的に時分割多重し、多重前の入力回線毎にメモリをパラレルに配置し、書き込み及び読み出しを各メモリでパラレルに行うことができる。

【0026】

また、前記共通スイッチ手段は、ビットスライス構成であり、スライスされた少なくとも1つのスイッチ手段を冗長に持つことにより、スライスされたスイッチ手段毎の保守・故障時の切り替えを可能とする。

【0027】

また、前記パケットバッファメモリ手段は、前記共通スイッチ手段の入力側及び出力側においてそれぞれ二重化され、出力側配置の前記パケットバッファメモリ手段へのデータ分配は予め前記共通スイッチ手段に設定した方路に基づいて行うことにより、保守・故障時の切り替えを可能とする。

【0028】

また、前記第1及び第2のポインタの制御手段と前記要求管理制御手段と前記スケジューリング処理手段とを含むスケジューラを分散配置し、かつ前記入力バッファメモリ手段を含む入力バッファ部の隣接する前記入力バッファ部間のスケジューリング情報を選択するためのスイッチ手段を更に設け、保守・故障時の切り替えを可能とする。

【0029】

本発明の他のパケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と；スケジューリング開始入力回線を示す第1のポインタの制御手段と；スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；前記パケットバッファメモリ手段は、バーストアクセス時のみ高速アクセス可能な大容量メモリを前段に、高速ランダムアクセス可能なメモリを後段に配し、通常は後段のメモリのみ使用し、後段メモリが一杯になったときは前段メモリに一時的に保持し、後段メモリに空きができた時点で移し替える。

【0030】

本発明の他のパケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と；スケジューリング開始入力回線を示す第1のポインタの制御手段と；スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、アクセスが低速な大容量メモリをパラレルに後段に配し、後段メモリへの書き込み動作は、前段メモリからパラレルに行い

、後段メモリの読み出し動作は競合が起こらないキューのみを選択して読み出しを行う。

【0031】

本発明の他のパケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と；スケジューリング開始入力回線を示す第1のポインタの制御手段と；スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；前記パケットバッファメモリ手段は、高速ランダムアクセス可能な小容量メモリを用いた待ち合わせキューを前段に、バーストアクセス時のみ高速アクセス可能な大容量メモリを用いた待ち合わせキューを後段に配し、後段メモリへの書き込み動作は、前段メモリに複数パケット溜まった時点でまとめて行い、後段メモリの読み出し動作は、複数パケットまとめて行う。

【0032】

本発明の他のパケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と；スケジューリング開始入力回線を示す第1のポインタの制御手段と；スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と；所望の出力回線に対する送出要求情報を保持する要求管理制御手段と；複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と；複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と；前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングする共通スイッチ手段と；前記パケットバッファメモリ手段のアドレス管理を行うアドレス管理手段とを備え；前

記パケットバッファメモリ手段は、複数の入力回線の固定長パケットを1つの入力回線上に確定的に時分割多重し、多重前の入力回線毎にメモリをパラレルに配置し、書き込み及び読み出しを各メモリでパラレルに行う。

【0033】

上記本発明の他のパケットスイッチ装置のそれぞれの構成において、前記共通スイッチ手段は、ビットスライス構成であり、スライスされた少なくとも1つのスイッチ手段を冗長に持つことにより、スライスされたスイッチ手段毎の保守・故障時の切り替えを可能とする。

【0034】

また、前記パケットバッファメモリ手段は、前記共通スイッチ手段の入力側及び出力側においてそれぞれ二重化され、出力側配置の前記パケットバッファメモリ手段へのデータ分配は予め前記共通スイッチ手段に設定した方路に基づいて行うことにより、保守・故障時の切り替えを可能とする。

【0035】

また、前記第1及び第2のポインタの制御手段と前記要求管理制御手段と前記スケジューリング処理手段とを含むスケジューラを分散配置し、かつ前記入力バッファメモリ手段を含む入力バッファ部の隣接する前記入力バッファ部間のスケジューリング情報を選択するためのスイッチ手段を更に設け、保守・故障時の切り替えを可能とする。

【0036】

本発明の別のパケットスイッチ装置は、ある一定速度でスケジューリング処理を行うスケジューリング処理手段と；入力回線速度分のスケジューリング速度で求められるパケットスロット時間を計測する第1のタイマ処理手段と；出力回線速度分のスケジューリング速度で求められるパケットスロット時間を計測する第2のタイマ処理部とを備え；ある入力回線に対するスケジューリング処理が行われたとき、前記第1のタイマ処理手段が計測を開始し、以降前記第1のタイマ処理手段が所定時間を計測する間は、その入力回線に対するスケジューリング処理を停止し、入力回線速度に応じたスケジューリングを実現する。

【0037】

この構成において、ある出力回線に対する確定が行われたとき、その出力回線に対する前記第2のタイマ処理手段の計測を開始し、以降前記第2のタイマ処理手段が所定時間を計測する間は、同一出力回線に対する確定を停止し、同一出力回線へのトラヒック流を出力回線速度以下に抑える。

【0038】

また、前記第1及び第2のタイマ処理手段の計測開始契機をスケジューリング時あるいは確定時とせずに、固定的な時間間隔とする。

さらに、パイプライン的にスケジューリング処理を行う際、スケジューリング速度と入力回線速度との比に基づく前後N段のパイプライン処理の該当入力回線に対するスケジューリングを停止し、入力回線速度に応じたスケジューリングを実現する。

【0039】

前記第2のタイマ処理手段は、パイプライン的にスケジューリング処理を行う際、パイpline番号と出力回線番号とにに基づきパイpline処理毎に独立に制御し、各パイpline処理内で所定の間隔を守ることで、平均的に各出力回線速度に応じたスケジューリングを実現する。

【0040】

【発明の実施の形態】

次に、本発明の実施の形態について図面を参照して説明する。

【パケットスイッチ装置のアドレス管理構成】

本発明の一実施の形態のパケットスイッチ装置において、メモリアドレスについては、アドレスをブロック単位で管理し、ブロック内の個別のアドレスは書き込み時または読み出し時にキュー毎に与えることでメモリ量を削減する。

【0041】

図12にブロック単位のアドレス管理手法を示す。アドレス管理メモリ12を構成する空きアドレスブロック管理メモリ（単に、空きアドレスFIFOと記載することもある）121及び論理出方路アドレス管理（読み出アドレスブロック管理）メモリ（単に、論理出方路または読み出アドレスFIFOと記載することもある）122は、それぞれブロック番号を管理する。また、論理出方路対応に、ブ

ロック内アドレスを示すシーケンスナンバー（SN）管理ポインタPNTを、書き込み・読み出し用にそれぞれ用意する。

【0042】

書き込み時は、該当する論理出方路キューの書込ブロックアドレスと書込シーケンスナンバー（SN）とにより書込アドレスを発行し、書込SNをインクリメントする。読み出し時は、該当する論理出方路キューの読出ブロックアドレスと読出シーケンスナンバー（SN）とにより読出アドレスを発行し、読出SNをインクリメントする。

【0043】

読み出し時または書き込み時において、読出SNまたは書込SNが一周したら、次のブロックアドレスを獲得する。これにより、例えば1ブロックを10パケットとすれば、アドレス管理メモリ12の容量を1/10に削減することができる。ここで、読出SN及び書込SNと読出ブロックアドレス及び書込ブロックアドレスとは出方路数に対応して用意すればよく、メモリ増加量は小さい。しかも、アドレスリンク方式を探る場合、そもそも出方路対応にスタートポインタ及びエンドポインタを用意する必要があり、この数とSN及びブロックアドレスのメモリ量とは同一であるため、メモリ量の増加は発生しない。

【0044】

アドレスリンク方式におけるマルチキャスト（MC）実現手法について図13を参照して説明する。マルチキャストは入力するマルチキャストパケットを一旦マルチキャストキューに格納し、それを各出方路の論理出方路キューに振り分けることで実現する。ここで、バッファメモリ（セルバッファメモリ131）のアクセス速度を抑えるため、パケット自体は振り分けを行わず、アドレスのみ各論理出方路アドレス FIFO132に振り分ける。論理出方路アドレス FIFO132はアドレスリンクにより構成されているため、異なる論理出方路のアドレス FIFOは同一のアドレスを使用できない。

【0045】

そこで、振り分け時には新たに空きアドレス FIFO133よりアドレスを取得することとし、そのアドレスと実際にパケットが格納されているアドレスとの

対応をとるためのバッファアドレステーブル134を用意し、振り分け時にこのテーブル134を設定しておく。読み出し時には、バッファアドレステーブル134で示されるアドレスよりパケットを読み出すことで、マルチキャストを実現することができる。

【0046】

ここで、バッファアドレステーブル134の同一バッファアドレスに対し、指定された全出方路へのマルチキャスト読み出しが終了しない限りは、読み出アドレスを空きアドレス FIFO133に返却できない。そこで、読み出し数を管理するためのマルチキャスト数管理テーブル（ビットマップ）135をバッファアドレス対応に用意し、全方路への読み出しが終了した時点で読み出アドレスを返却する。この手法はアドレスリンクを適用しない場合も同様である。

【0047】

次に、アドレスをブロック管理するときのマルチキャスト手法について図14を参照して説明する。基本的には、図13を参照して説明したマルチキャスト手法と同様であるが、アドレスがブロック管理されているため、振り分け時のアドレスは書込ブロックアドレスと書込SNとから取得する。また、ブロックアドレスの空きアドレス FIFO133への返却は、同一ブロックに含まれるアドレスのマルチキャストが全て終了しないと行うことができない。そこで、ブロックアドレス毎に、各SNの読み出しが終了したかどうかを管理するためのマルチキャスト数管理テーブル136を持ち、全SNへの読み出しが終わった時点でブロックアドレスを返却する。これにより、ブロックアドレス管理でのマルチキャストを実現することができる。

【0048】

〔パケット（セル）バッファメモリ構成〕

パケット（セル）バッファメモリを構成するために、大容量のDRAMをバッカアップとして使用した例を図15に示す。このセルバッファメモリ15は通常、小容量のSRAMで構成する入力バッファメモリ151を用い、入力回線IHWからのセル（パケット）のバッファリングを行うが、SRAMが一杯になった場合、一時的にDRAMで構成する入力バッファメモリ152にバーストアクセ

スで入力回線 I HW からのセルデータを保持する。SRAM にバースト長分の空きが生じた時点で、DRAM から SRAM に動作を移し替える。これにより、高速アクセスを実現しつつ大容量のセルバッファメモリ 15 を構成することができる。

【0049】

また、セルバッファメモリを構成するために、SRAM の入力バッファメモリを入力回線 I HW からのセルの一時保持用として使用した例を図 16 に示す。入力回線 I HW からのセルは、論理出方路キュー毎に SRAM で構成する一時保持用の入力バッファメモリ 161 に格納され、キュー毎にバースト数分溜まった時点で大容量のDRAM で構成する入力バッファメモリ 162 にセルデータを移し替える。DRAM からの読み出しは、バースト数分（複数パケット（セル）分）まとめて行う。これにより、高速アクセスを実現しつつ大容量のセルバッファメモリ 16 を構成することができる。

【0050】

ここで、負荷の低いケースでは、前段の SRAM にバースト数分のセルデータが溜まるのに時間がかかる可能性がある。これに対し、タイマを設け規定時間を超えたら溜まる前に後段のDRAM に書き込んでしまう方法と、前段の SRAM もスケジューリング対象とし、その時点で溜まっているセルデータを読み出してしまう方法により対処可能である。

【0051】

セルバッファメモリを構成するために、複数のDRAM をパラレルに配置した例を図 17 に示す。使用するDRAM は論理出方路対応で分け、例えば全部で 128 個の回線対応の出方路の場合は、回線 0 - 31 対応の出方路は論理出方路キュー RAM 0、回線 32 - 63 対応の出方路は論理出方路キュー RAM 1、回線 64 - 95 対応の出方路は論理出方路キュー RAM 2、及び回線 96 - 127 対応の出方路は論理出方路キュー RAM 3 に割り当てる。

【0052】

入力バッファメモリ 172 を構成するDRAM の前段には、一時保持用の入力バッファメモリ 171 を構成する SRAM が同じ割り当てでパラレルに用意され

る。入力したセルは、論理出方路に対応するSRAMに書き込まれ、一つずつパラレル配置のDRAMに移し替えられる。DRAMからの読み出しは、スケジューリング時に分割RAM間でぶつからないように割り当てる。つまり、ある回線が選択されたら、その後3回は、スケジューラがその回線と同一のRAMを使用している回線を対象としない。これにより、高速アクセスを実現しつつ大容量セルバッファメモリ17を構成することができる。

【0053】

セルバッファメモリを構成するために、出力回線OHW上に確定的な位置にセルを時分割多重した場合の例を図18に示す。この構成は図17に示した構成に類似しているが、あらかじめ確定的に分離されているため、パラレル化したDRAM182に対応するセルが連續で到着することはない。よって、前段配置のSRAMが不要になる。その代わり、パラレル化された各DRAM182から任意の出方路に出力されるため、それぞれのDRAM182を論理出方路キューに分割する必要がある。これにより、高速アクセスを実現しつつ大容量セルバッファメモリ18を構成することができる。

【0054】

〔冗長構成〕

スイッチ部のコアスイッチの冗長構成例を図19に示す。スイッチ部191のコアスイッチ1911についてN+1ビットスライスで構成した場合、同一のスライススイッチ1911が複数（N個=16）並列に配置されることになる。このスライススイッチ1911を1つあるいは複数余計に用意することにより、切り替えが可能となる。切り替えは、入力バッファ部192の任意のビットにおいてN個の現用系スライススイッチから冗長（予備系）スライススイッチ1912に出力することを選択するセレクタ1921と、入力バッファ部192の受け側で各ビットが冗長スライススイッチ1912からのデータかを選択するセレクタ1922, 1923とを用意することで可能となる。

【0055】

また、それぞれセレクタを任意のビットから任意のスライススイッチに切り替えできるようにすると、どのスライススイッチが冗長スライススイッチかを意識

することなく切り替えることも可能である。

【0056】

入力バッファ部のパケット（セル）バッファメモリ202の冗長構成を図20に示す。送信回線側及び受信回線側については対向の二重化装置との関係で二重化したいケースもあるが、スイッチ部201のコアスイッチがN+1冗長構成であるため、交絡を設けられない。そこで、コアスイッチにコピー機能を設けておき、現用（ACT）系からの入力セルのみを、スイッチ部201に通すことで、二重化することができる。スイッチ部201のコピー方法は、セルにコピー制御ビットを付与する方法と、スイッチ部自体に設定する方法とがある。

【0057】

スケジューリング情報の切り替え構成を図21に示す。入力バッファ部のパッケージ化されたパケットバッファメモリ211を抜いてもスケジューリング情報の受け渡しができるように、パケットバッファメモリ211に対応にバイパススイッチ（スケジューリング情報切替部）212を設ける。そして、バイパスするときは、パケット（セル）バッファメモリ211における制御遅延分を加味することにより、途中で動作を止めることなく切り替えることが可能になる。

【0058】

図21におけるバイパススイッチ212の詳細構成例を示す図22を参照すると、バイパススイッチ212には隣接バイパススイッチ及び入力バッファ部からの2系統のスケジューリング信号が入力される。それぞれのスケジューリング信号の同期信号を同期信号抽出部2120, 2121で抽出し、図示省略の制御部からの基準フレーム信号との位相差を位相差検出部2122, 2123で検出する。その結果を位相調整バッファ2124, 2125の読み出制御部2126, 2127に通知し、位相差を吸収する。

【0059】

また、タイミング調整部2128において基準フレーム信号に同期した上記制御部からの切替信号によって、出力スケジューリング信号の切替えをセレクタ2129により行う。スケジューリング信号の同期信号としては、フレーム信号の並走や周期的な同期パターン（例えば、固定パターンやCRC）などがある。書

制御部2130, 2131は位相調整バッファ2124, 2125へのスケジューリング信号の書き込みを制御する。

【0060】

図23に上記バイパススイッチ212の動作例を示す。

(a) 初期状態では、基準フレーム信号に従いセレクタ2129を通してループを構成する。

(b) 入力バッファ部を構成する新しいパッケージが組み込まれると、まず入力バッファ部においてスケジューリング信号のフレーム検出により、同期確立を行う。

(c) 次に、バイパススイッチ212において、入力バッファ部側についてスケジューリング信号のフレーム検出により、同期確立を行う。

(d) 同期確立後、バイパススイッチ212のセレクタ2129を切り替えることにより、増設した入力バッファ部のスケジューリング信号が送出される。

(e) 入力バッファ部を構成する新しいパッケージを更に増設するときは、同様の手順を踏む。

(f) パッケージの取り替えは、バイパススイッチ212のセレクタ2129の切替後に行う。

【0061】

〔第1のスケジューラの構成・動作〕

本発明の一実施の形態のパケットスイッチ装置における入力バッファ部のスケジューラの構成を示す図24を参照すると、要求管理部221は入力バッファメモリから通知される送出要求数を管理する。第1のタイマ処理部(1)222は自己入力回線IHWに対するスケジューリング周期を計測する。第2のタイマ処理部(2)223は各論理出方路に対応に備えられ、その出力回線に対するスケジューリング割り当ての周期を計測する。スケジューリング処理部224はタイマ処理部(1)222でスケジューリング可能と判断されたときにスケジューリング処理を行い、要求管理部221に送出要求のある出力回線で、かつタイマ処理部(2)223で割り当て可能と指示された出方路のなかからスケジューリングアルゴリズムに従って出方路を確定するものである。各スケジューラ22は入力

回線（入力方路）IHW対応に設けられる。

【0062】

この構成において、ある入力回線に対するスケジューリング処理を行ったとき、タイマ処理部（1）222において、所定時間（スケジューリング速度／入力回線速度）の計測を開始する。そして、所定時間が経過するまでは、スケジューリング処理部224に対してスケジューリング処理不可の停止指示を行って入力回線速度以上の速度でスケジューリングを行わないようとする。したがって、スケジューラ22から入力バッファメモリに対して送出要求を通知する契機も入力回線速度となり、その結果、低速回線インターフェースには高速な入力バッファメモリが不要となる（動作例1）。

【0063】

また、ある入力回線のスケジューリング処理過程において、ある出力回線（出方路）OHWが確定されたとき、タイマ処理部（2）223において、所定時間（スケジューリング速度／出力回線速度）の計測を開始する。そして、所定時間が経過するまでは、スケジューリング処理部224に対して、その出力回線に対する割り当て不可の停止指示を行って、出力回線速度以上でスケジューリングが行われないようにしている。したがって、スケジューラ22が入力バッファメモリに対して、その出力回線に送出せよという指示の周期も、出力回線速度以下となり、その結果、出力回線に対してその回線速度以上のトラヒック流入を抑止でき、大容量の入力バッファメモリが不要となる（動作例2）。

【0064】

スケジューリング周期が1単位時間では完結せず、数単位時間かけてパイプライン処理を行う際に、上記動作例1で説明した動作を行うために、あるパイプライン処理段のある入力回線のスケジューリング処理が行われたとき、そのパイプライン処理段の前後N（スケジューリング速度／入力回線速度）段のスケジューラに対して、その入力回線に対するスケジューリング処理不可を指示する。この結果、スケジューラから入力バッファメモリに対して行われる読み出し指示が入力回線速度以上で行われることを抑制できる（動作例3）。

【0065】

また、スケジューリング周期が1単位時間では完結せず、数単位時間かけてパイプライン処理を行う際に、上記動作例2で説明した動作を行うために、あるパイプライン処理段のある入力回線のスケジューリング処理において、ある出力回線が確定されたとき、そのパイプライン段のスケジューラ22に対して、N（スケジューリング速度／出力回線速度）パイプライン周期の間、その出力回線に対する割り当て不可を指示する。この結果、出力回線に対してその回線速度以上のトラヒック流が流入しなくなり、大容量の入力バッファメモリが不要となる（動作例4）。

【0066】

タイマ処理部（1）222及びタイマ処理部（2）223の計測契機を周期的に行うようにし、タイマ処理部222、223の負荷軽減を図る。この結果、上記動作例2及び4に比べると、出力トラヒック流のバースト性が強まる傾向にあるが、タイマ処理部の負荷軽減を図るとともに、周期的にスケジューリングまたは割り当て許可の指示が現れるため、他のスケジューラにどの出力回線が確定不可であるかを通知する必要がなく、スケジューラ間の情報受け渡しの量を削減することができる（動作例5）。

【0067】

〔パケットスイッチ装置の動作〕

本発明の一実施の形態のパケットスイッチ装置の動作について、図25から図37を順次参照して説明する。図25は書き込み動作手順、図26、図27に書き込み動作例、図28にマルチキャスト動作手順、図29から図32にマルチキャスト動作例、図33に読み出し動作手順、図34から図37に読み出し動作例を示す。

【0068】

（ポインタ基本処理）

まず、アドレス FIFOへの追加動作及びアドレス FIFOからの読み出動作について説明する。

【0069】

（1）アドレス FIFOへの追加動作：

アドレス FIFO が空のときは、追加するアドレスをスタート (START) ポインタ S-PNT 及びエンド (END) ポインタ E-PNT にそれぞれ設定する。一方、アドレス FIFO が空で無いときは、END ポインタが示すアドレスリンクに追加するアドレスを設定する。更に、END ポインタに追加するアドレスを設定する。

【0070】

(2) アドレス FIFO から読出動作 :

START ポインタが示すアドレスを獲得する。START ポインタが示すアドレスリンクより、NEXT アドレスを獲得する。更に、START ポインタに NEXT アドレスを設定する。

【0071】

アドレス FIFO が空かどうかは、キュー長カウンタより識別する。ただし、空きアドレス FIFO (空きアドレスブロック管理メモリ) はキュー長カウンタを持たない。そこで、空きアドレス FIFO は必ず空きにならないように、次のように管理する。つまり、初期はアドレス 0 番を空きアドレス FIFO に割り当てる。空きアドレス生成カウンタはアドレス 1 番から書込アドレスを生成する。書込アドレス獲得時、START ポインタと END ポインタとが等しいならば、入力セルを廃棄する。

【0072】

(書き込み動作)

到着セルをセルバッファメモリに書き込み、対応する論理出方路キューのアドレスリンクを更新する。書込アドレスは空きアドレス FIFO の先頭より獲得する。ただし、初期状態では空きアドレス FIFO には何も設定されていない。そこで、初期動作用に書込アドレスを生成するカウンタ、つまり初期アドレス生成カウンタINI-CNT を設ける。初期アドレス生成カウンタINI-CNT が最大 (MAX) 値になるまでは、このカウンタより書き込みアドレスを獲得し、MAX 値になったらカウンタを止め、以降は空きアドレス FIFO より書き込みアドレスを獲得する。

【0073】

空きアドレスFIFOから書込アドレスを獲得できなかった場合は、入力セルを廃棄する。また、マルチキャスト(MC)動作との整合のため、バッファアドレス領域にも書込アドレスを保持する。更に、MCルーティングビット領域に対し、ユニキャストセルの場合は出方路番号をビットマップに変換した値を、マルチキャストセルの場合はMCルーチングビットを保持する。

【0074】

書き込み手順は次に示すS1からS7である。つまり、

S1：書込アドレスの獲得、

S2：到着セルを書込アドレスに対応するセルバッファメモリに書き込む、

S3：到着セルのヘッダ情報に対応する論理キューのアドレスFIFOに書込アドレスを追加、

S4：書込アドレスをバッファアドレス領域に保持、

S5：入力セルがユニキャストセルの場合、出方路番号に対応するMCルーティングビットをセット、

S6：入力セルがマルチキャストセルの場合、MCルーティングビットを保持、

S7：対応する論理キュー、共通バッファ（セルバッファメモリ）のキュー長カウンタをインクリメントする。

【0075】

(マルチキャスト動作)

論理マルチキャストキューのアドレスFIFOよりMCアドレスを読み出し、これをMC先頭アドレスとしてレジスタに保持する。次に、対応するMCルーティングビットをSSRAMより読み出し、MC残ルーティングビットとしてレジスタに保持する。そして、MC残ルーティングビットを参照し、対応する論理出方路キューにアドレスを振り分ける。振り分けは、1パケット（セル）時間に1出方路づつ行う。振り分けるアドレスは、最初の出方路はMC先頭アドレスを、2回目以降は新たに空きアドレスFIFOから取得したMC追加アドレスを使用する。

【0076】

ただし、実際のセルはMC先頭アドレスに対応するセルバッファメモリに書き

込まれているため、バッファアドレスとしてMC先頭アドレスを新たに取得したアドレス毎に保持する。マルチキャストの1セル目であるかどうかは、MC先頭アドレス有効レジスタにより判断する。

【0077】

マルチキャスト処理手順は次に示すS11からS18及びS21からS26である。つまり、

1セル目：

S11：論理マルチキャストキューのMCアドレスをMC先頭アドレスレジスタに保持、

S12：MC先頭アドレスに対応するMCルーティングビットをMC残ルーティングビットレジスタに保持、

S13：MC残ルーティングビットが立っている回線のうち、一番若い回線を選択、

S14：選択回線に対応する論理キューのアドレスFIFOにMC先頭アドレスを追加、

S15：MC先頭アドレスに対応するバッファアドレス領域にMC先頭アドレスを設定、

S16：MC残ルーティングビットに対し、選択回線に対応するビットを「0」に更新、

S17：対応するMCキューのキュー長カウンタをデクリメント、

S18：対応する論理キューのキュー長カウンタをインクリメントする。

2セル目以降：

S21：空きアドレスFIFOよりMC追加アドレスを獲得、

S22：MC残ルーティングビットが立っている回線のうち、一番若番の回線を選択、

S23：選択回線に対応する論理キューのアドレスFIFOにMC追加アドレスを追加、

S24：MC追加アドレスに対応するバッファアドレス領域にMC先頭アドレスを設定、

S 2 5 : MC 残ルーティングビットに対し、選択回線に対応するビットを「0」に更新、

S 2 6 : 対応する論理キュー、共通バッファのキュー長カウンタをインクリメントする。

【0078】

MC 残ルーティングビットがオール「0」になるまでこの動作を続ける。オール「0」になったら、そのセルのマルチキャストは終了し、MC 先頭アドレス有効レジスタをリセットする。次のスロットで新たなMC アドレスを論理マルチキャストキューより獲得し、始めの動作に戻る。また、空きアドレス FIFO よりアドレスが獲得できなかった場合は、そのスロットでのマルチキャスト動作を止め、アドレスが獲得できるまで待つ。

【0079】

(読み出し動作)

対応する論理出方路キューのアドレス FIFO より読出アドレスを獲得し、読出アドレスに対応するバッファアドレスのセルバッファメモリより、セル読み出しを行う。その後、読出アドレスがバッファアドレスに一致しないときのみ、読出アドレスの返却を行う。さらに、バッファアドレスに対応するMC ビットマップの読出回線に対応するビットをクリアする。その結果、オール「0」であれば、対応するセルの読み出しが全て完了したと判定し、バッファアドレスを空きアドレス FIFO に返却する。

【0080】

読み出し手順は次に示す S 3 1 から S 3 6 である。つまり、

S 3 1 : 論理出方路キューのアドレス FIFO より、読出アドレスを獲得、

S 3 2 : 読出アドレスに対応するバッファアドレスより、セルを読み出す、

S 3 3 : バッファアドレスと読出アドレスとが一致しない場合、読出アドレスをアドレス FIFO に追加し、共通バッファのキュー長カウンタをデクリメント、

S 3 4 : バッファアドレスに対応するMC ビットマップに対し、読出回線に対応するビットをクリア、

S 3 5 : クリア後のMC ビットマップがオール「0」なら、バッファアドレスを

アドレス FIFO に追加し、共通バッファのキュー長カウンタをデクリメント、
S36：対応する論理キューのキュー長カウンタをデクリメントする。

【0081】

〔第2のスケジューラの構成・動作〕

本発明の一実施の形態のパケットスイッチ装置における入力バッファ部のスケジューラの構成を示す図38を参照すると、要求管理部361は入力バッファメモリから通知される送出要求数を管理する。第1のタイマ処理部(1)362は自己入力回線に対するスケジューリング周期を計測する。第2のタイマ処理部(2)363は各論理出方路に対応に備えられ、その出力回線に対するスケジューリング割り当ての周期を計測する。スケジューリング処理部364はタイマ処理部(1)362でスケジューリング可能と判断されたときにスケジューリング処理を行い、要求管理部361に送出要求のある出力回線で、かつタイマ処理部(2)363で割り当て可能と指示された出方路のなかからスケジューリングアルゴリズムに従って出方路を確定するものである。

【0082】

各スケジューラ(SCH)36は入力回線(入力方路)IHWに対応に設けられ、1単位時間内、ここでは、1パケット(セル)時間内、各入力回線で出力方路が競合しない様に、自己入力回線のパケット送出出方路を決定すべく、スケジューリングアルゴリズムに従って出力回線を確定する。

【0083】

(入力回線速度スケジューリングの必要性)

図39は9.6 Gbpsの伝送速度のポートを有するN×Nコアスイッチ(スイッチ部)371に、伝送速度の異なる入力回線を入力バッファ部372、373、374を通して収容した例を示す。この例では、伝送速度2.4 Gbpsのインターフェースの入力回線#0及び伝送速度4.8 Gbpsのインターフェースの入力回線#5が伝送速度の異なる入力回線として収容されている。

【0084】

この入力回線の収容形態において、何も速度制御を行わなければ、各スケジューラ(SCH)3721、3731は単位時間(ここでは、9.6 Gbps)の

速度で、各入力回線に対してスケジューリングを行ってしまい、2. 4 G b p s のバッファリング能力しか持たない入力回線#0や4. 8 G b p s のバッファリング能力しかない入力回線#5に対して、9. 6 G b p s の速度で送出指示を送ってしまう。すなわち、入力バッファ部372, 373に高速な入力バッファメモリ3722, 3732が必要になる。

【0085】

(入力回線速度でスケジューリングを行う動作)

図40を用いて入力回線速度でスケジューリングを行う例を説明する。入力バッファ部に高速バッファメモリを設置することを回避するためには、各入力回線は下記のパケット(セル)時間間隔以上でスケジューリングされなければいけない。

入力回線#0 : 4パケット時間 (= 9. 6 G b p s / 2. 4 G b p s)

入力回線#5 : 2パケット時間 (= 9. 6 G b p s / 4. 8 G b p s)

時刻T=1において、入力回線#0に対するスケジューリング処理が行われたとする。ただし、スケジューリングの結果、実際に割り当てが行われたか否かは問わない。このとき、スケジューラ#0のタイマ処理部(1)362(図39参照)は、所定間隔の計測を開始し、スケジューリング処理部364に対してスケジューリング停止信号を通知する。所定時間の計測は、パケット時間毎にカウンタのインクリメントを行う単純なカウンタ回路で実現できる。そして、スケジューリング停止信号は、スケジューリング処理が行われたときにセットされ、所定の計測時間経過した時点で解除される。

【0086】

時刻T=2では、まだ所定の時間が経過していない、つまりスケジューリング停止信号がL₀w(0)レベルであるため、スケジューリングを行うことができない。このように、スケジューリング処理部364は、スケジューリング停止信号をスケジューリング処理のイネーブル信号として用いており、停止信号がL₀wレベルの時はスケジューリング処理を行わない。

【0087】

入力回線#0のスケジューリング停止信号は、時刻T=5において解除される

、つまりHigh(1)レベルになるため、時刻T=5では入力回線#0のスケジューリングを行うことができる。上記制御によって、入力回線#0のスケジューリング処理周期は、4パケット時間間隔となる。図40中、スケジューリング処理が行われるタイムスロットを斜線で示している。

【0088】

入力回線#5の処理も上記と同様であり、時刻T=1において、入力回線#5のスケジューリング処理が行われたとすると、スケジューラ#5のタイマ処理部(1)362が所定時間の計測を開始して時刻T=2経過までスケジューリング停止信号を送出し、スケジューラ364は、時刻T=2においては入力回線#5のスケジューリングを行わない。

【0089】

上記の様に、スケジューリング処理が行われた時点から所定時間の間、スケジューリング停止信号をタイマ処理部(1)362からスケジューリング処理部364に通知することで、スケジューリング契機が入力回線速度周期で行われるため、必然的にスケジューラ36から入力バッファ部への読み出し指示の契機も入力回線速度となる。この結果、入力バッファ部には、その入力回線速度と同じ速度で読み出し可能なバッファメモリだけがあればよい。

【0090】

入力回線速度に応じたスケジューリングの処理手順S391からS398を図41に示す。S392からS395はスケジューリング処理プロセスである。単位時間毎にイベントが発生すると(S391)、S392において、スケジューリング停止信号の論理で現在のスケジューリング状態を判定し、判定結果が「1」の場合、すなわち、スケジューリング可能状態であればS393からS395の処理を行う。一方、判定結果が「0」の場合(スケジューリング不可)には、スケジューリング処理を行わない。

【0091】

上記判定結果、「1」のケースでは、S393で所定の入力回線のスケジューリング処理を行い、S394及びS395で、スケジューリング停止信号及びスケジューリング停止区間を管理するタイマ(タイマ1)を「0」にセットする。

【0092】

S396からS398は、スケジューリング停止信号の更新処理である。パケット時間ごとにタイマをインクリメントし(S396)、その結果が所定の時間を超えているか否かを判定する(S397)。タイマ値が所定時間を超えていた場合には、スケジューリング停止信号を「1」にセットする(S398)。

【0093】

このように、タイマ処理によって、所定時間(=1／入力回線速度)の間はスケジューリング停止信号が「0」にセットされる。この結果、所定時間の間は、スケジューリング処理が実行されないため、入力回線以上の速度でスケジューリング処理を行わないように制御することができる。

【0094】

(出力回線速度スケジューリングの必要性)

図42には伝送速度2.4Gbpsの出力回線#Cに対して、伝送速度2.4Gbpsの入力回線#Aと伝送速度4.8Gbpsの入力回線#Bからのパケットをコアスイッチ401を通して送出した例を示している。この場合、入力回線#A対応の入力バッファ部402からはコアスイッチ401に伝送速度2.4Gbpsでパケットが送出され、かつ入力回線#B対応の入力バッファ部403からはコアスイッチ401に伝送速度4.8Gbpsでパケットが送出されるため、出力回線#C対応の出力バッファ部404の出力バッファメモリ4041には、合計7.2Gbpsのトラヒックが流入することになる。

【0095】

この7.2Gbpsのトラヒック流を受信する出力回線#C対応の出力バッファ部404は、2.4Gbpsの読み出し速度しかもっておらず、パケット廃棄を回避するためには大容量の出力バッファメモリ4041が必要となる。

【0096】

(出力回線速度でスケジューリング処理の確定を行う動作)

図43を参照して出力回線速度に応じたスケジューリング動作を説明する。出力側のバッファメモリ量を小さくするためには、その出力回線に流入するトラヒックを出力回線速度以下に抑えなければならない。したがって、全入力回線にお

いてひとつの出力回線に送出する間隔をその出力回線速度に制御すべく、この例では、各出力回線へのスケジューリング割り当ての間隔を以下の時間で設ける必要がある。

出力回線#2：2パケット時間（=9.6 Gbps / 4.8 Gbps）

出力回線#4：4パケット時間（=9.6 Gbps / 2.4 Gbps）

時刻T=2において、入力回線#1のスケジューリング処理が行われ出力回線#2が確定したとする。この時、入力回線#1のスケジューラは、他の全スケジューラに対して出力回線#2を確定したことを探知する。そして、この通知を受信した各スケジューラは、自己タイマ処理部（1）の出力回線（この例では、出力回線#2）に対応するタイマの計測を開始し、自己スケジューリング処理部に対して割り当て停止信号を通知する。タイマ処理部（1）と同様に、タイマ処理部（2）もパケット時間毎にインクリメントを行う、カウンタ回路を用いて所定時間の計測を行うことができる。

【0097】

割り当て停止信号は、スケジューリング割り当てが行われるか、他スケジューラからの確定通知を受信した際にセットされ、タイマ処理部（2）の所定時間経過した後に解除される。スケジューリング処理部は、この割り当て停止信号が通知されていない（すなわち停止状態でない）出力回線に対して送出確定を行う。

【0098】

この例の時刻T=3では、出力回線#2の割り当て停止信号（Lowレベル）が通知されているため、どのスケジューラも出力回線#2に対する割り当ては行えない。そして、時刻T=4において、タイマ処理部（2）の計測が完了し、出力回線#2の割り当て停止信号が解除されている。その結果、時刻T=4の入力回線#5において、あらためて出力回線#2への割り当てが行われている。図4-3中、出力回線#2への割り当てを斜線で示している。

【0099】

一方、出力回線#4への割り当ても同様の手順で処理が行われる。時刻T=2において入力回線#5のスケジューリング処理が行われ、出力回線#4の確定が行われたとき、同様の手順で、全スケジューラに対して出力回線#4の確定を行

ったことを通知する。前述した様に各スケジューラのタイマ処理部（2）の出力回線#4に対する所定時間の計測を開始し、出力回線#4の割り当て停止信号をスケジューリング処理部に通知する。図43中、出力回線#4に対する割り当てをクロス斜線で示す。

【0100】

このように、この例においては、時刻T=3では、出力回線#2及び出力回線#4の割り当てを行うことができず、また時刻T=4～T=5では、出力回線#4の割り当てが行われない。

【0101】

同図中の斜線部とクロス斜線部とをそれぞれ個別に見てみると、出力回線#2への割り当て契機（図中、斜線ブロック）は、少なくとも2パケット時間以上あいており、また出力回線#4への割り当ては、4パケット以上あいている。この時間間隔は、それぞれの出力回線にパケットが到着することと等価であり、各出力方路へのトラヒックを出力回線速度以下に抑えることができる事を示している。

【0102】

出力回線速度に応じたスケジューリングの処理手順S4201からS4214を図44に示す。S4202からS4204は割り当て停止タイマ処理の初期化シーケンスである。単位時間毎にイベントが発生すると（S4201）、S4202において、他スケジューラから確定通知を受信したときに、自スケジューラの割り当て停止信号を「0」にセットし（S4203）、タイマ（タイマ2）を「0」にセットする（S4204）。確定通知を受信しなかった場合は、何も処理を行わない。

【0103】

S4205からS4208はスケジューリング処理及びスケジューリング処理後の割り当て停止タイマ処理の初期化シーケンスである。S4205では、自スケジューラの割り当て停止信号#N_M（N：入力回線番号、M：出力回線番号）が「1」にセットされている出力回線の中から送出回線を確定するようスケジューリング処理を行う。そして、スケジューリング処理が確定した後、確定し

た出力回線に対応する割り当て停止信号#N_M (N:確定入力回線番号、M:確定出力回線番号)、及び対応するタイマ (N:確定入力回線番号、M:確定出力回線番号) を「0」にセットする (S4207, S4208)。

【0104】

S4209からS4213は次時刻のスケジューリング停止信号を更新するシーケンスである。S4209において、入力回線数 (K) 分の処理が終了したかを判定し、全入力回線の更新処理が終了した後、出力回線番号mを「0」にセット (S4214) して処理を終了する。本図では、処理をシリーズに行っているが、並列処理で行っても良い。S4210は、入力回線数 (K) 分の処を行ったか否かをカウントするカウンタであり、上記並列処理の際には不要となる。

【0105】

S4211からS4213は、スケジューリング停止信号の更新シーケンスである。S4211において、パケット時間毎にタイマ値をインクリメントし (S4211)、その結果が所定時間 (=1/出力回線レート) を超えているかを判定する (S4212)。S4212において、所定時間を超過していると判定された場合には、対応する割り当て停止信号を「1」にセットする (S4213)。

【0106】

このように、S4211からS4213において、所定時間 (=1/出力回線レート) の間、割り当て停止信号を「0」にセットして、S4205において各々のスケジューラが、各出力回線に対して各出力回線レート以上で割り当てを行わないように、割り当て停止信号が「0」の出力回線に対して割り当てを行わないように制御している。

【0107】

〔第3のスケジューラの構成・動作〕

本発明の一実施の形態のパケットスイッチ装置における入力バッファ部のスケジューラの構成を示す図45を参照すると、このスケジューラ43は各パイプライン処理ごとにスケジューリング処理部434とタイマ処理部(2)433とをそれぞれ有している。要求管理部431及びタイマ処理部(1)432は図38

に示す第2のスケジューラと同様に動作する。

【0108】

次に、1回のスケジューリング処理をMパケット時間かけて、パイプライン的にスケジューリングを行う場合の回線速度に応じたスケジューリング方式について説明する。

【0109】

図46にパイプラインシーケンスを示す。同図中、各時刻(T)対応の四角の中の数字は、スケジューリング対象の入力回線番号を示す。ここでは、 4×4 スイッチを想定して、4パケット時間かけて1回のスケジューリングを完了させる例をとって説明する。4パケット時間費やしてスケジューリング処理を行い、パケット時間毎に結果を得るために、4倍の並列度を持つパイプライン構成となる。

【0110】

各スケジューラ43は、自分の入力回線番号の位置でスケジューリング処理を行う。例えば、時刻T=1では、入力回線#0のパイプライン処理#0、入力回線#1のパイプライン処理#0、入力回線#3のパイプライン処理#2、及び入力回線#2のパイプライン処理#2のスケジューリング処理が行われる。同図中の太線TTは、スケジューラ43から入力バッファメモリに送出指示を通知することを表しており、時刻T=1～T=3のパイプライン処理#0のスケジューリング結果が時刻T=4で対応の入力バッファメモリに通知されることを意味する。また、空欄の時刻にはスケジューリング処理を行わないことを示している。

【0111】

(入力回線速度に応じたスケジューリング割り当ての方法)

図46に示す上記パイプライン構成において、例えば、時刻T=8かつパイプライン処理#3において、入力回線#3のスケジューリング処理が行われたとする。入力回線#3は4.8 G b p sの回線速度であるため、スケジューリング間隔は2パケット時間(スケジューリング速度/入力回線速度 = 9.6 G b p s / 4.8 G b p s)に制限しなければならない。ここで確定した結果は、実際には、時刻T=11で入力バッファメモリに通知されるため、時刻T=10及びT=

12では、入力回線#3に対する読出指示を与えないようにする必要がある。

【0112】

そのため、時刻T=8でスケジューリング処理が行われた後、そのパイプライン処理段の前後N段（ここで、Nはスケジューリング速度／入力回線速度で求められ、この例では1）のパイプライン処理内で入力回線#3に対するスケジューリングを行わないように、入力回線#3のスケジューリング処理を行ったとき、前後1段、すなわちパイpline#2とパイpline#0に対して、入力回線#3に対するスケジューリング不可信号を通知する。同図中、入力回線#3に対するスケジューリング不可の区間を斜線で表記した。

【0113】

同様に入力回線#0の場合も、その回線がスケジューリングされた時点から前後N段（N=3）のパイpline処理段に対してスケジューリング停止信号を通知し、入力回線#0へのスケジューリングを行わないように制御する。

【0114】

このように、パイpline構成を用いる場合でも、そのパイpline段の前後N段のスケジューリング処理部434にスケジューリング停止信号を通知することで、入力回線速度に応じたスケジューリングが可能となる。

【0115】

（入力回線速度でスケジューリングを行う動作）

入力回線速度に応じたスケジューリング動作の処理手順S451からS456を図47に示す。同図中、条件1は、停止信号を送出するスケジューリング周期を示している。例えば、図46において、時刻T=20で伝送速度622Mbpsの入力回線#1のスケジューリングが行われたとする。この場合、上述したようにスケジューリング確定点から前後16段（9.6Gbps/622Mbps）のパイpline処理の間、スケジューリング停止を行わなければならない。

【0116】

そこで、各パイpline処理部（タイマ処理部）が、それぞれ幾つ分のパイpline処理時間待たなければならないかの停止パイpline周期を下記の式で求める。

【0117】

停止パイプライン周期 = $(F_{st} PPL + K - PP) / PPLN$

ここで、 $F_{st} PPL$: スケジューリング処理を行ったパイプライン番号

K : スケジューリング速度／入力回線速度

PP : 自己パイプライン番号

$PPLN$: 総パイプライン段数

この例における停止パイプライン周期はそれぞれ次のようになる。

パイプライン#0 : $(1 + 16 - 0) / 4 = 4.25 = 4$ 周期

パイプライン#1 : $(1 + 16 - 1) / 4 = 4.00 = 4$ 周期

パイプライン#2 : $(1 + 16 - 2) / 4 = 3.75 = 3$ 周期

パイプライン#3 : $(1 + 16 - 3) / 4 = 3.50 = 3$ 周期

(出力回線速度に応じたスケジューリング割り当ての方法)

パイプライン構成を探る場合は、各パイプライン処理ごとに独立に出力速度を守るように制御する。図4.8を参照して出力回線速度に応じたスケジューリング処理を説明する。例えば、時刻 $T = 8$ かつパイプライン処理 #3 の入力回線 #2 のスケジューリング結果が出力回線 #0 に確定されたとき、そのパイプライン周期の次の N パイプライン周期の間、パイプライン処理 #3 では出力回線 #0 に対する割り当てを行わないようとする。出力回線 #2 の回線速度が 2.4 Gbps のため、上記 N 周期は次式で求められる。

【0118】

$N = \text{スケジューリング速度} / \text{出力回線速度} (N = 9.6 \text{ Gbps} / 2.4 \text{ Gbps} = 4)$

時刻 $T = 10$ において、パイプライン処理 #3 かつ入力回線 #1 のスケジューリング処理時、出力回線 #3 が確定したときも同様の手段で次の 1 周期の間、パイプライン処理 #3 では出力回線 #3 に対する割り当てを行わない。ここで、時刻 $T = 9$ でも、パイプライン処理 #0 かつ入力回線 #0 のスケジューリング処理において、同一出力回線 #3 への割り当てが行われており、その結果、時刻 $T = 11$ 及び $T = 12$ で連続して出力回線 #3 へパケットが流入する。しかし、これは長い周期でみれば、平均的に出力回線速度のレートで送出されている。

【0119】

このように、パイプライン構成では、各パイプライン処理毎に所定の間隔で割り当てを行うことで、平均的にその出力回線速度以下の速度でパケットが流入するため、出力バッファメモリ量は少なくて済む。

【0120】

図49に出力回線速度に応じたスケジューリング動作の処理手順S471からS479を示す。S472からS474はスケジューリング処理及びスケジューリング確定後の割り当て停止信号、タイマの初期化シーケンスである。単位時間毎にイベントが発生すると（S471）、S472では、現時刻のスケジューリング対象の入力回線について、割り当て停止信号が「1」にセットされている出力回線の中から送出回線を確定する。そして、確定したパイプライン処理段、入力回線、出力回線に対応する割り当て停止信号、及びタイマ値を「0」にセットする（S473, S474）。

【0121】

S475からS479は割り当て停止信号の更新プロセスである。S475において、全パイプライン処理段分の処理が終了したかを判定し、全パイプライン処理段分の処理が完了した後、処理を終了する。S476は、パイプライン処理の終りであるかを判定するプロセスであり、通常、N段のパイプライン処理段があると、ひとつのパイプライン処理はN単位時間で完結する。このプロセスは各パイプライン処理の最終時刻を判定するものであり、最終時刻の判定は各パイプライン処理毎にカウンタを持って、カウンタ値がNになった時に最終時刻と判定しても良いし、外部からパイプライン処理の最終時刻である旨の指示を与えて良い。

【0122】

S477からS479は割り当て停止信号の更新プロセスであり、S476において、パイプライン処理の最終時刻に、対応するタイマ値をインクリメントし（S477）、その結果が所定の周期（=1／出力回線レート）を超えているかを判定する（S478）。所定の周期を超えている場合には、対応する割り当て停止信号を「1」にセットする。

【0123】

このように、S477からS479において、所定の周期の間、割り当て停止信号を「0」にセットして、各パイプライン処理段のスケジューリング処理S472において、各出力回線に対して各出力回線レート以上で割り当てを行わないよう、割り当て停止信号が「0」の出力回線に対して割り当てを行わないよう制御し、この制御を各パイプライン処理段毎に独立に行うことによって、平均的に所定の出力回線レート以上の割り当てを行わないよう制御することが可能となる。

【0124】

〔第4のスケジューラの構成・動作〕

上記第3のスケジューラにおいては、スケジューリング割り当てを契機として、タイマ処理部（2）のタイマ計測を開始しており、パケットが読み出されてから計測を行うため、出力回線の入り口でみたとき、バースト性の低いトラヒックとなる。しかし、同一パイプライン処理内の異なるスケジューラに対して、あとどれくらい停止信号を送出すべきかを伝達する必要があり、隣接スケジューラに対して情報通知を行う必要がある。

【0125】

次に述べる第4のスケジューラでは、この情報通知を削減すべく、全スケジューラで共通のフレーム信号を持ち、そのフレーム信号に従って出力回線の割り当てを行う。そのフレーム信号の形式を図50に示している。なお、第4のスケジューラの構成は図45に示す第3のスケジューラの構成と同一である。

【0126】

(出力回線速度に応じたスケジューリング処理)

各スケジューラ内にパイプライン処理ごとに1パケット時間づつ位相がずれている出力回線速度に応じた固定フレーム信号を有している。図50中のフレーム信号FRはパイプライン処理#3のものである。この信号のHighレベルの領域でその出力回線に対するスケジューリング割り当てが可能な領域とし、Lowレベルの区間を停止区間とすることで、前述のスケジューラ間の情報伝達が不要になる。同図中、各パイプライン処理のフレーム信号のLowレベルの領域を斜

線で示している。しかし、同図からも分かるように、ひとつの出力回線に連続してN個（N：パイプライン数）割り当てが行えるため、N個連続してパケットが送出され、トラヒック流のバースト性が強くなる。

【0127】

【変形例】

上記実施の形態では、パイpline処理内の入力回線番号の並びは一例であり、どのような並びであってもよい。また、上記実施の形態では、入力回線速度に応じたスケジューリングと出力回線速度に応じたスケジューリングとの動作について、それぞれ個別に説明しているが、入力回線速度に応じてスケジューリング対象入力回線を決定し、出力回線速度に応じてスケジューリング割り当てを確定することで、入出力ともに速度対応のスケジューリングを行うことができる。

【0128】

さらに、上記実施の形態では、入力バッファ型スイッチ装置について詳述したが、出力バッファ部に入力バッファ部と同様の構成を採用した入出力バッファ型スイッチ装置（図51参照）にもこの発明は適用できる。

【0129】

【発明の効果】

以上説明したように、本発明によれば、メモリアドレスをブロック単位で管理し、ブロック内の個別のアドレスは書き込み時または読み出し時にキー毎に与えることでメモリ量を削減することができる。

【0130】

また、本発明によれば、少数の回線収容時でも必要となっていたパケット多重部（MUX）や低速インターフェースにおける高速バッファメモリの配置を不要とし、低速から高速まで全ての回線インターフェースを混在収容することができる。

【0131】

また、本発明によれば、高速アクセスを実現しつつ大容量のパケット（セル）バッファメモリを構成することができる。

【0132】

さらに、本発明によれば、出力側のパケット多重分離部（DMUX）における

所要バッファメモリ量を削減することができる。

【図面の簡単な説明】

【図1】 パケットスイッチ装置の入力バッファ部を説明するためのブロック図。

【図2】 パケットスイッチ装置の入力バッファ部を説明するためのブロック図。

【図3】 パケットスイッチ装置の入力バッファ部及びスイッチ部を説明するためのブロック図。

【図4】 パケットスイッチ装置の入力バッファ部を説明するためのブロック図。

【図5】 パケットスイッチ装置のアドレス管理FIFOを説明するための図。

【図6】 パケットスイッチ装置のDRAMのバーストアクセスを説明するための図。

【図7】 パケットスイッチ装置の入力バッファ部及びスイッチ部の二重化構成を説明するためのブロック図。

【図8】 パケットスイッチ装置のパケットバッファメモリとアドレスリンクとの共用を説明するためのブロック図。

【図9】 パケットスイッチ装置のマルチキャスト時のアドレス管理FIFOを説明するための図。

【図10】 パケットスイッチ装置のDRAMのパラレルアクセスを説明するための図。

【図11】 パケットスイッチ装置のスケジューリング情報の切断を説明するための図。

【図12】 本発明のパケットスイッチ装置のアドレスブロック管理手法を説明するための図。

【図13】 アドレスリンクにおけるマルチキャストを説明するための図。

【図14】 ブロック管理手法におけるマルチキャストを説明するための図。

【図15】 セルバッファメモリの構成例を示すブロック図。

- 【図16】セルバッファメモリの構成例を示すブロック図。
- 【図17】セルバッファメモリの構成例を示すブロック図。
- 【図18】セルバッファメモリの構成例を示すブロック図。
- 【図19】コアスイッチの冗長構成例を示すブロック図。
- 【図20】セルバッファメモリの冗長構成例を示すブロック図。
- 【図21】スケジューリング情報の切替構成を示すブロック図。
- 【図22】バイパススイッチの詳細構成を示すブロック図。
- 【図23】バイパススイッチの動作例を示すブロック図。
- 【図24】第1のスケジューラの構成を示すブロック図。
- 【図25】書き込み動作を説明するための図。
- 【図26】ユニキャスト書き込み動作を説明するための図。
- 【図27】マルチキャスト書き込み動作を説明するための図。
- 【図28】マルチキャスト動作を説明するための図。
- 【図29】マルチキャスト動作を説明するための図。
- 【図30】マルチキャスト振り分け動作を説明するための図。
- 【図31】マルチキャスト振り分け動作を説明するための図。
- 【図32】マルチキャスト振り分け動作を説明するための図。
- 【図33】読み出し動作を説明するための図。
- 【図34】読み出し動作を説明するための図。
- 【図35】読み出し動作を説明するための図。
- 【図36】読み出し動作を説明するための図。
- 【図37】読み出し動作を説明するための図。
- 【図38】第2のスケジューラの構成を示すブロック図。
- 【図39】入力バッファ部の処理能力の問題を説明するための図。
- 【図40】入力回線速度に応じたスケジューリング処理を説明するための図
- 。
- 【図41】入力回線速度に応じたスケジューリング処理手順を説明するための図。
- 【図42】出力トラヒック流の問題を説明するための図。

【図4 3】 出力回線速度に応じたスケジューリング処理を説明するための図

【図4 4】 出力回線速度に応じたスケジューリング処理手順を説明するための図。

【図4 5】 第3のスケジューラの構成を示すブロック図。

【図4 6】 入力回線速度に応じたスケジューリング処理を説明するための図

【図4 7】 入力回線速度に応じたスケジューリング処理手順を説明するための図。

【図4 8】 出力回線速度に応じたスケジューリング処理を説明するための図

【図4 9】 出力回線速度に応じたスケジューリング処理手順を説明するための図。

【図5 0】 第4のスケジューラにおける出力回線速度に応じたスケジューリング処理を説明するための図。

【図5 1】 入出力バッファ型スイッチ装置の構成例を示すブロック図。

【符号の説明】

1 2 アドレス管理メモリ

1 2 1 空きアドレスブロック管理メモリ

1 2 2 読出アドレスブロック管理メモリ

1 3 1, 1 5, 1 6, 1 7, 1 8, 1 9 2, 2 0 2, 2 1 1 セルバッフ
アメモリ

1 3 2 アドレスFIFO

1 3 3 空きアドレスFIFO

1 9 1, 2 0 1 スイッチ部

2 2, 3 6, 4 3 スケジューラ

2 2 1, 3 6 1, 4 3 1 要求管理部

2 2 2, 3 6 2, 4 3 2 タイマ処理部(1)

2 2 3, 3 6 3, 4 3 3 タイマ処理部(2)

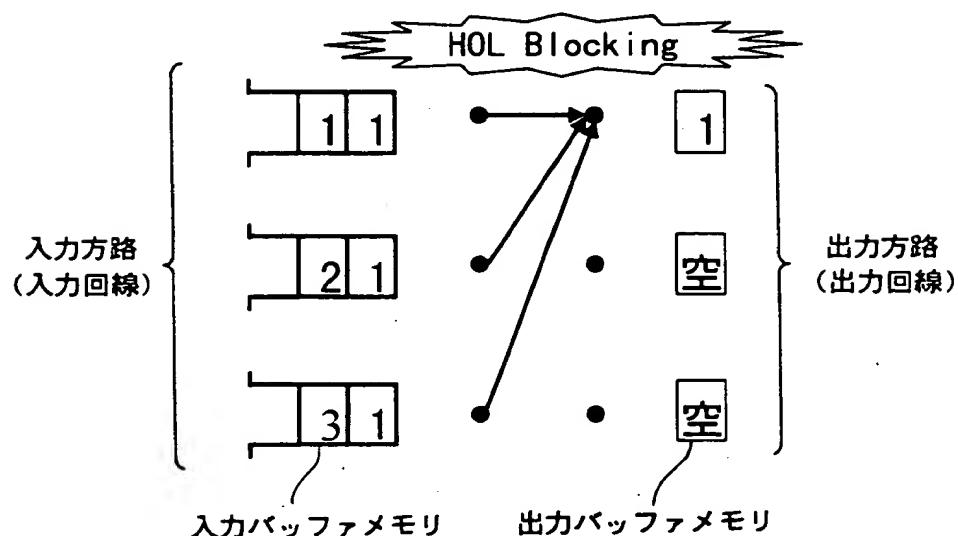
特平11-235596

224, 364, 434 スケジューリング処理部

【書類名】 図面

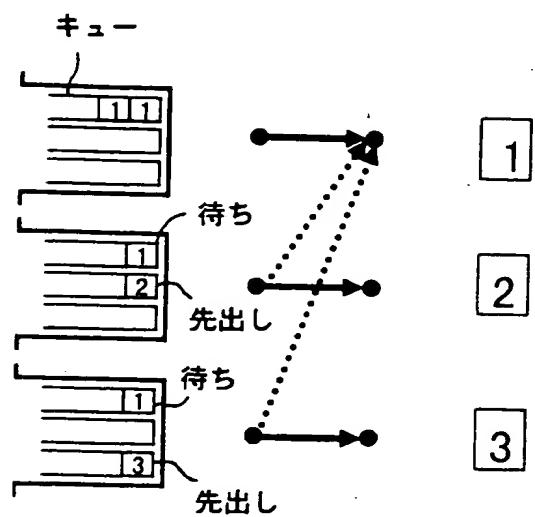
【図1】

パケットスイッチ装置の入力バッファ部を
説明するためのブロック図



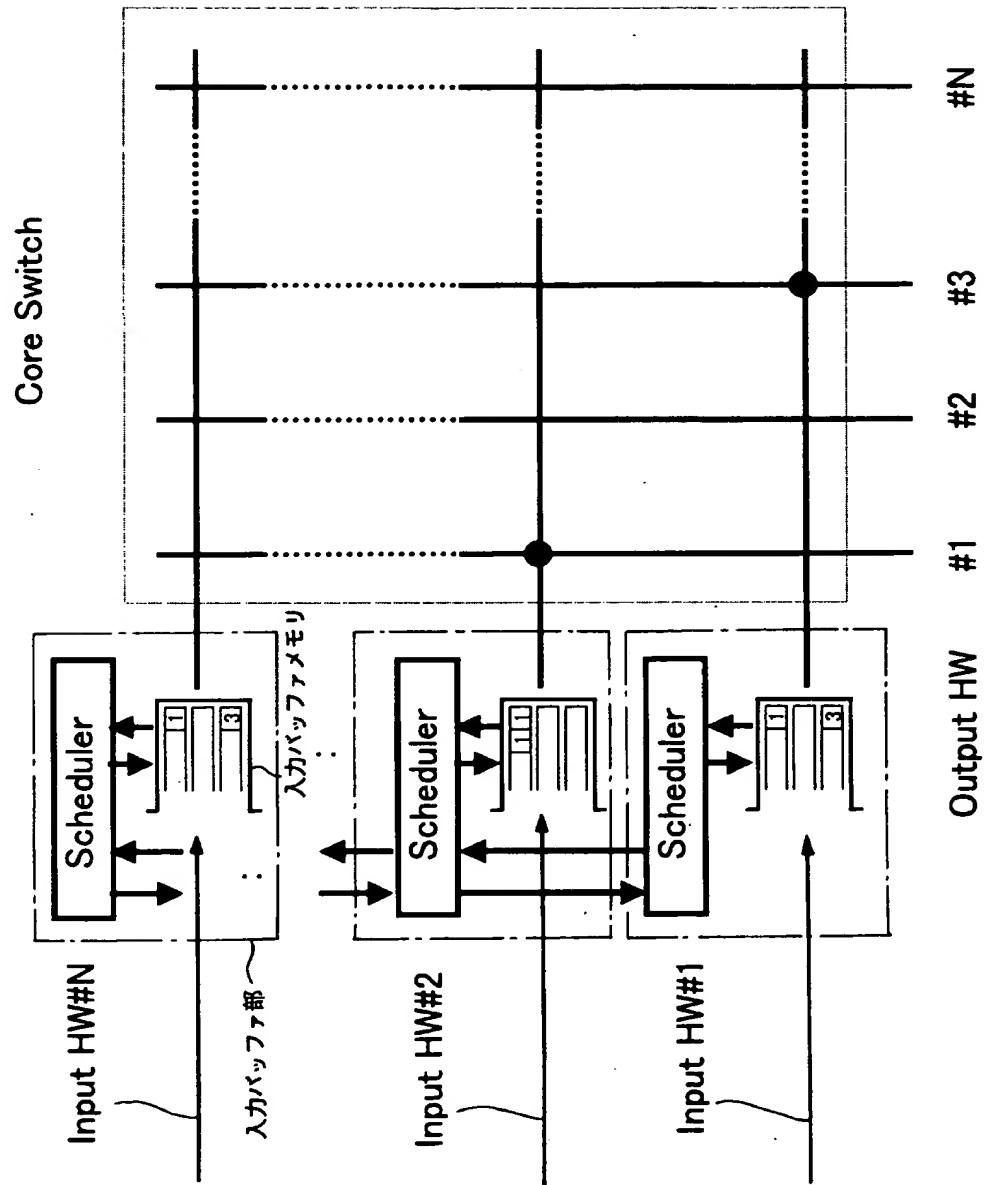
【図2】

パケットスイッチ装置の入力バッファ部を
説明するためのブロック図



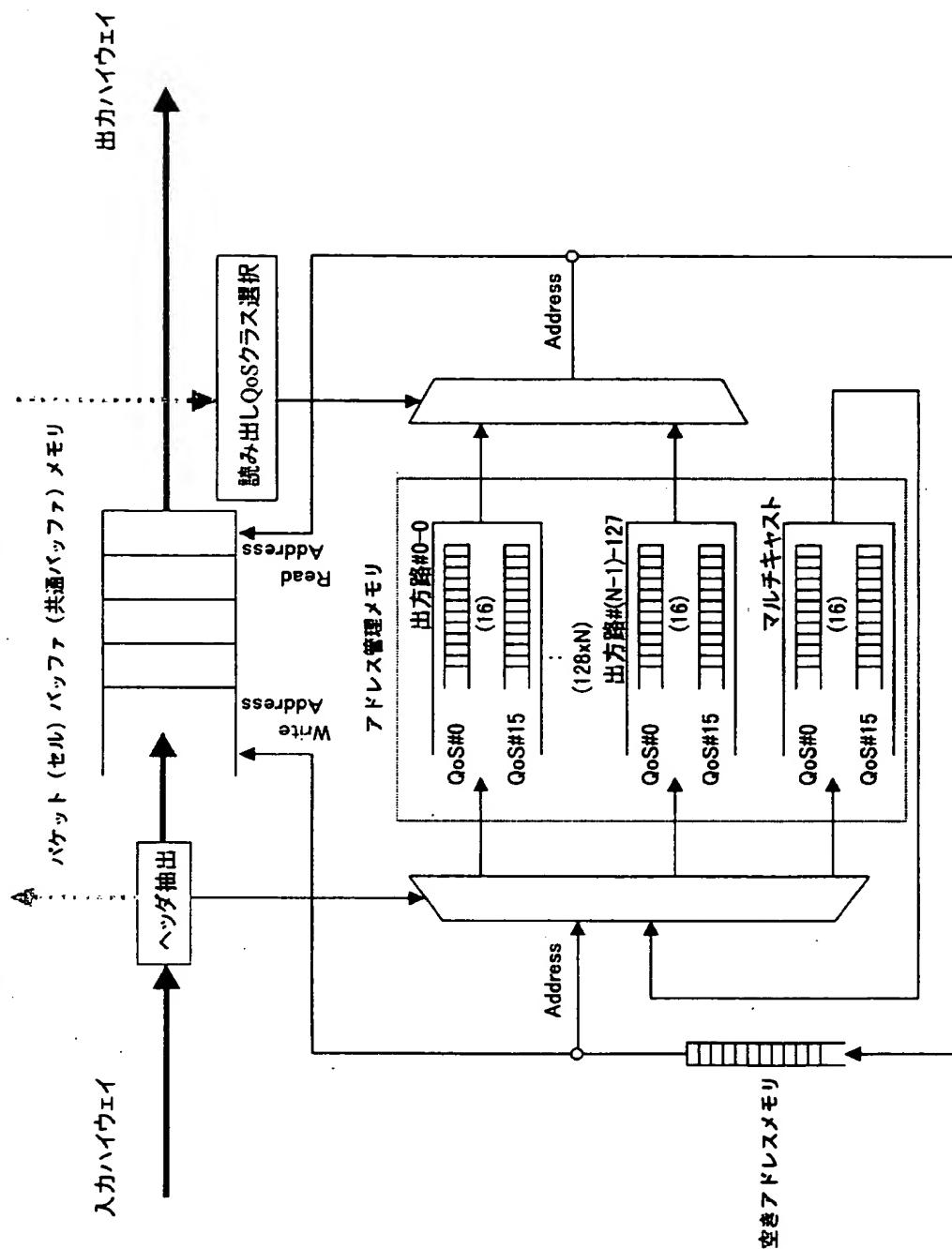
【図3】

パケットスイッチ装置の入力バッファ部及びスイッチ部を説明するためのブロック図



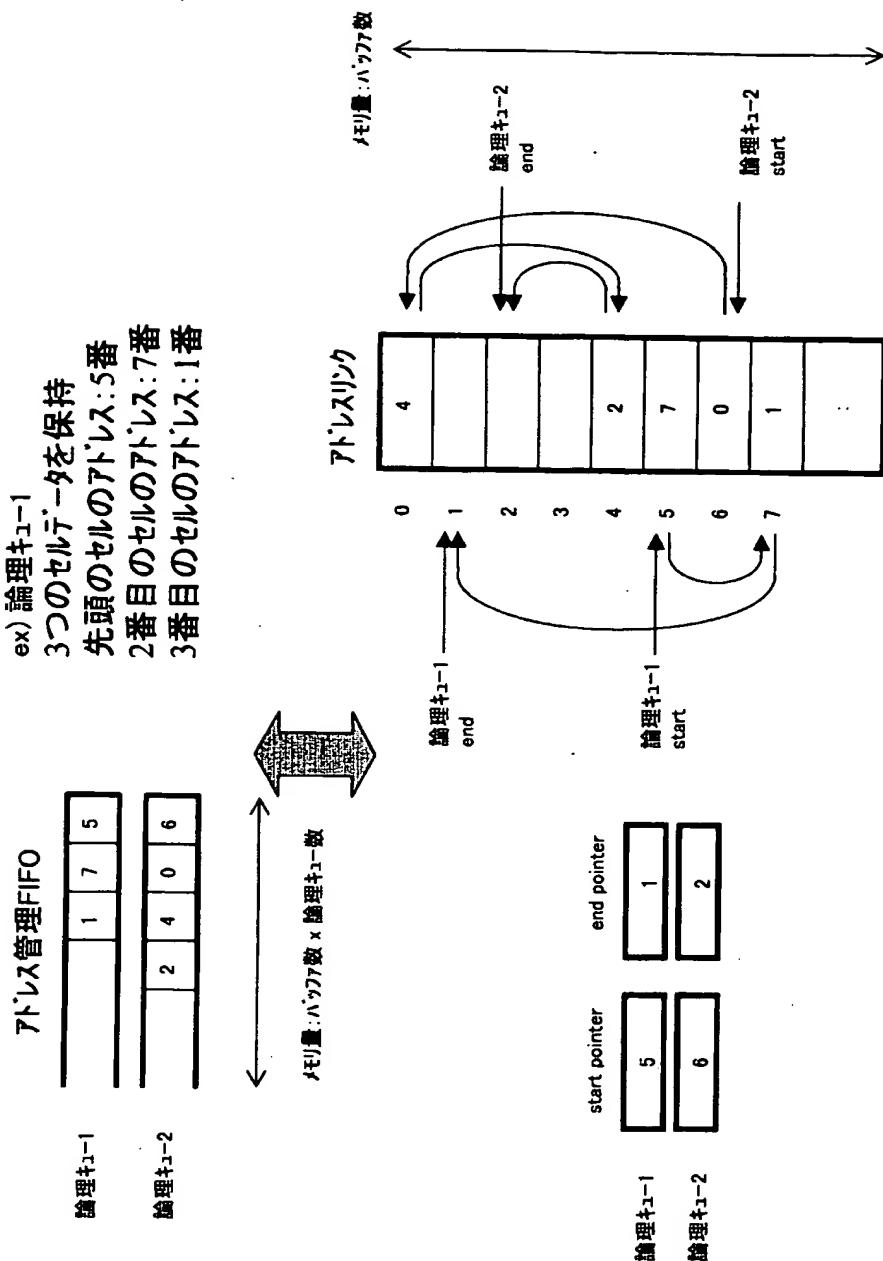
【図4】

到着回線＆クラス通知(スケジューラへ)
ヘッダ世出
読み出し回線指示(スケジューラより)



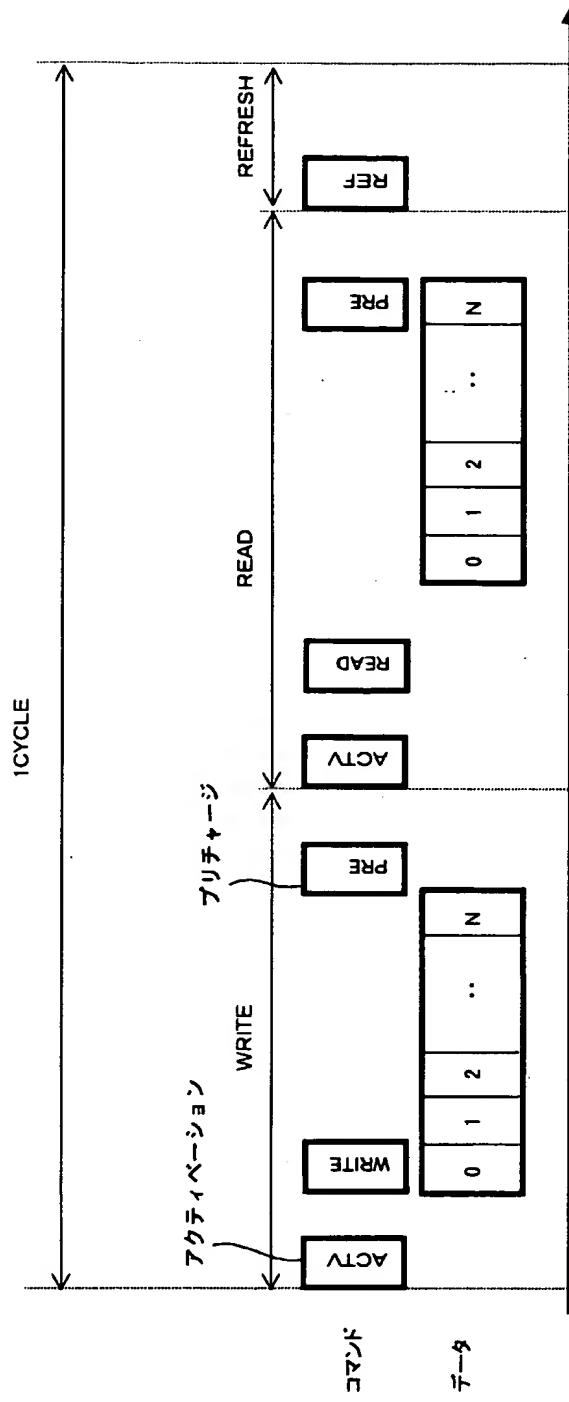
【図5】

パケットスイッチ装置のアドレス管理 FIFOを説明するための図



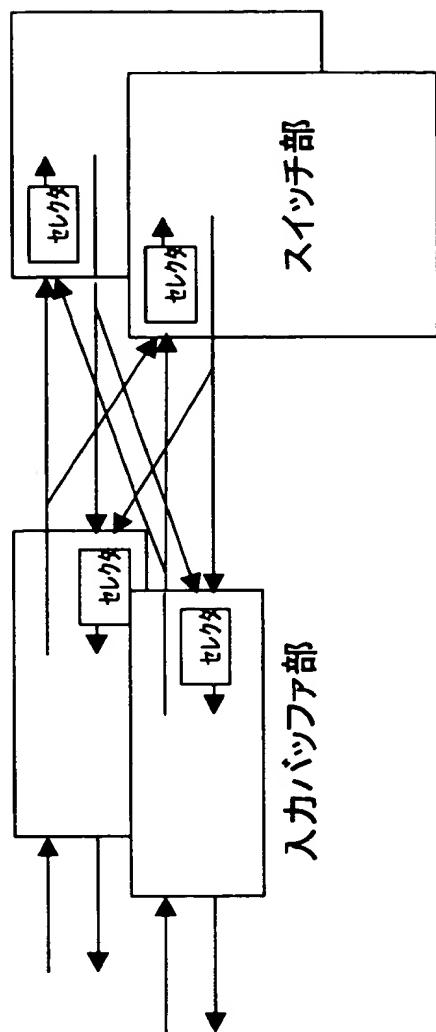
【図6】

パケットスイッチ装置のDRAMのバーストアクセスを説明するための図



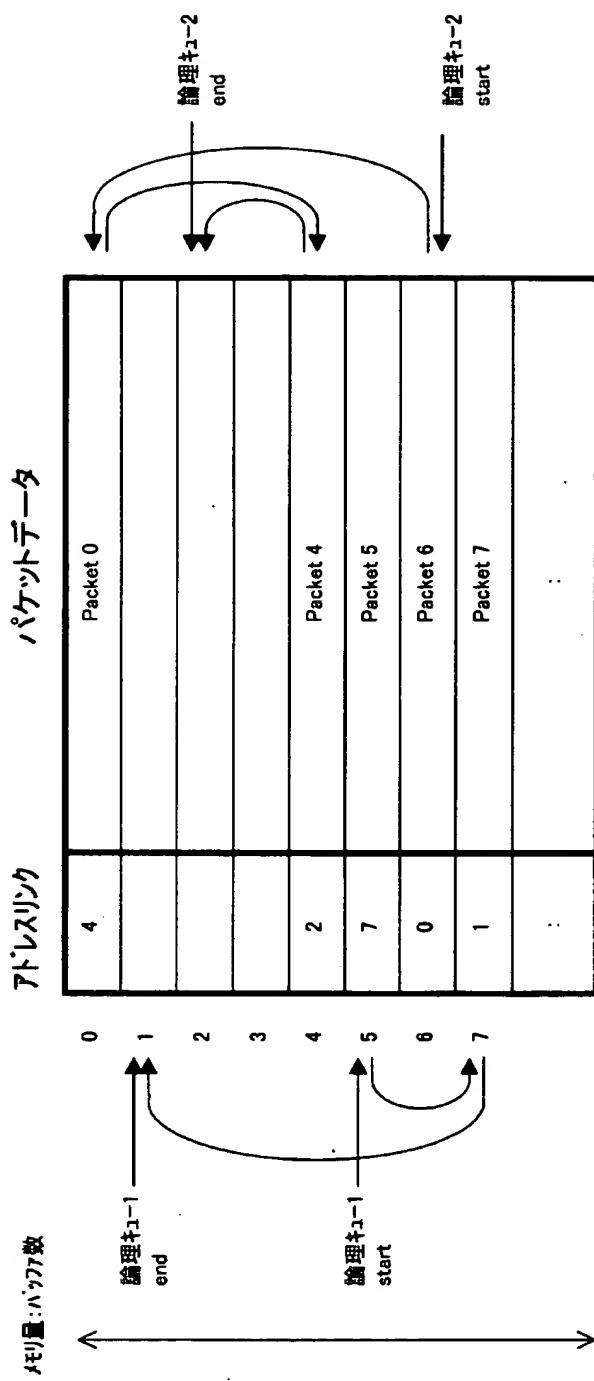
【図7】

パケットストレーナ装置の入力バッファ部及びスイッチ部の二重化構成を説明するためのブロック図



【図8】

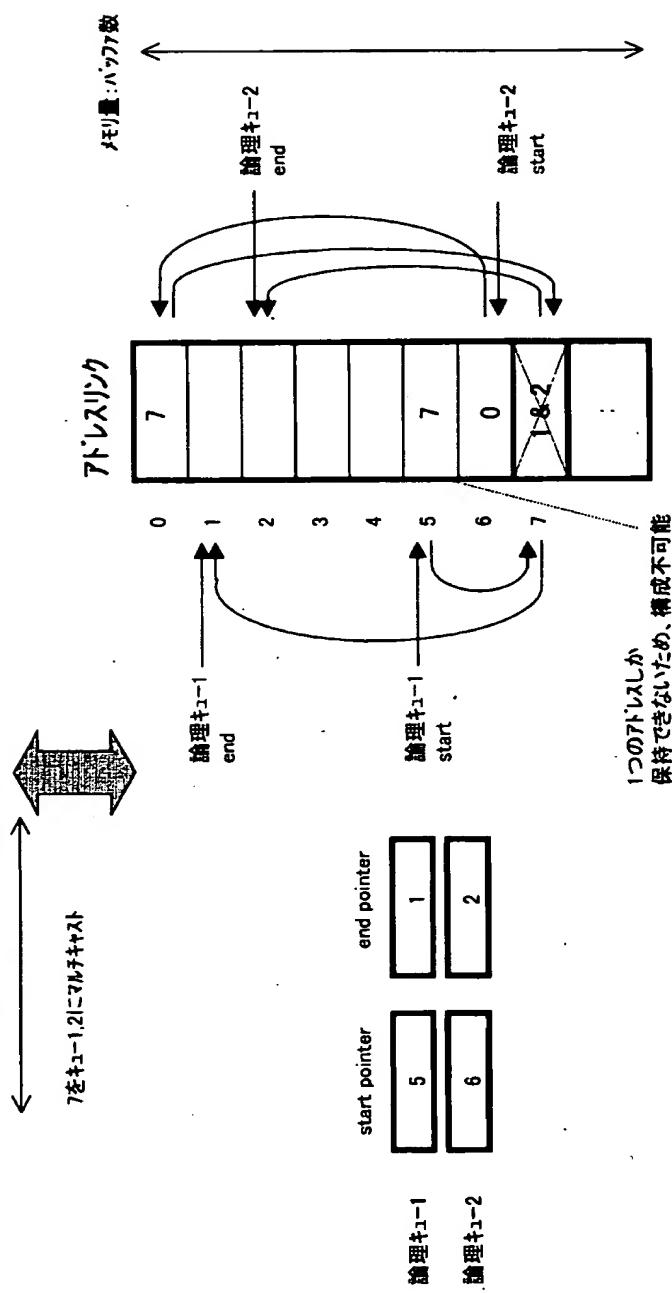
パケットスイッチ装置のパケットバッファメモリとアドレスリンク
との共用を説明するためのブロック図



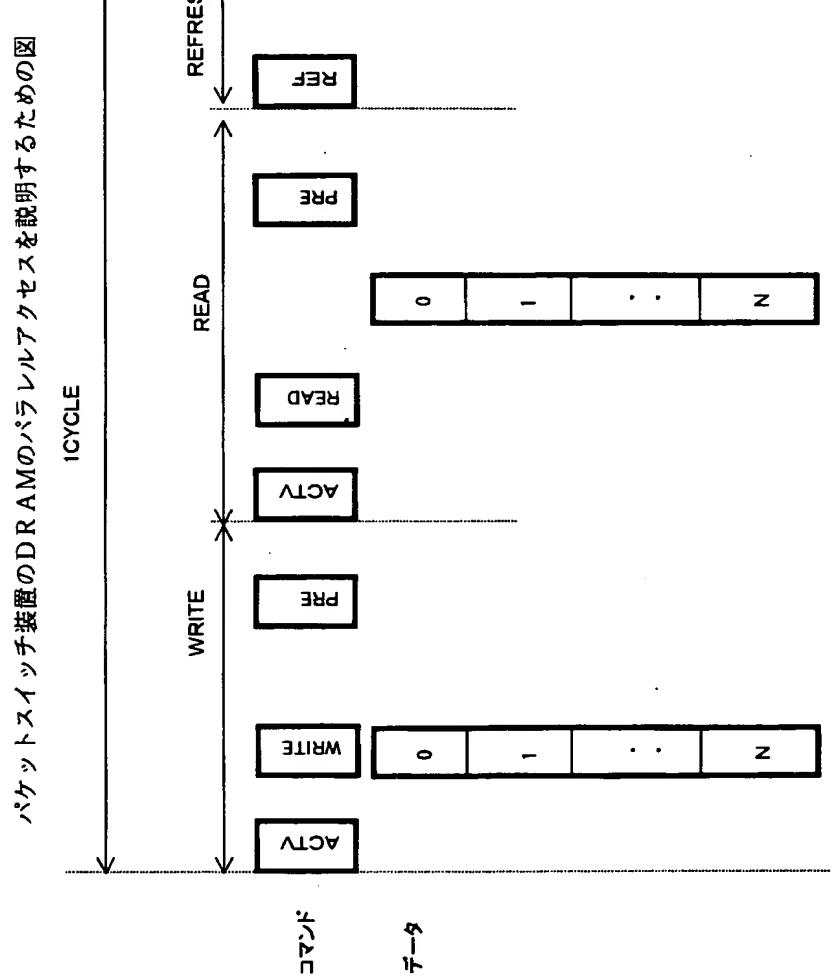
【図9】



論理キューライン	アドレス管理FIFO	ex) 論理キューライン				
論理キューライン	<table border="1"><tr><td></td><td>1</td><td>7</td><td>5</td></tr></table>		1	7	5	3つのセルデータを保持
	1	7	5			
論理キューライン	<table border="1"><tr><td>2</td><td>7</td><td>0</td><td>6</td></tr></table>	2	7	0	6	先頭のセルのアドレス: 5: 2番目のセルのアドレス: 1: 3番目のセルのアドレス:
2	7	0	6			

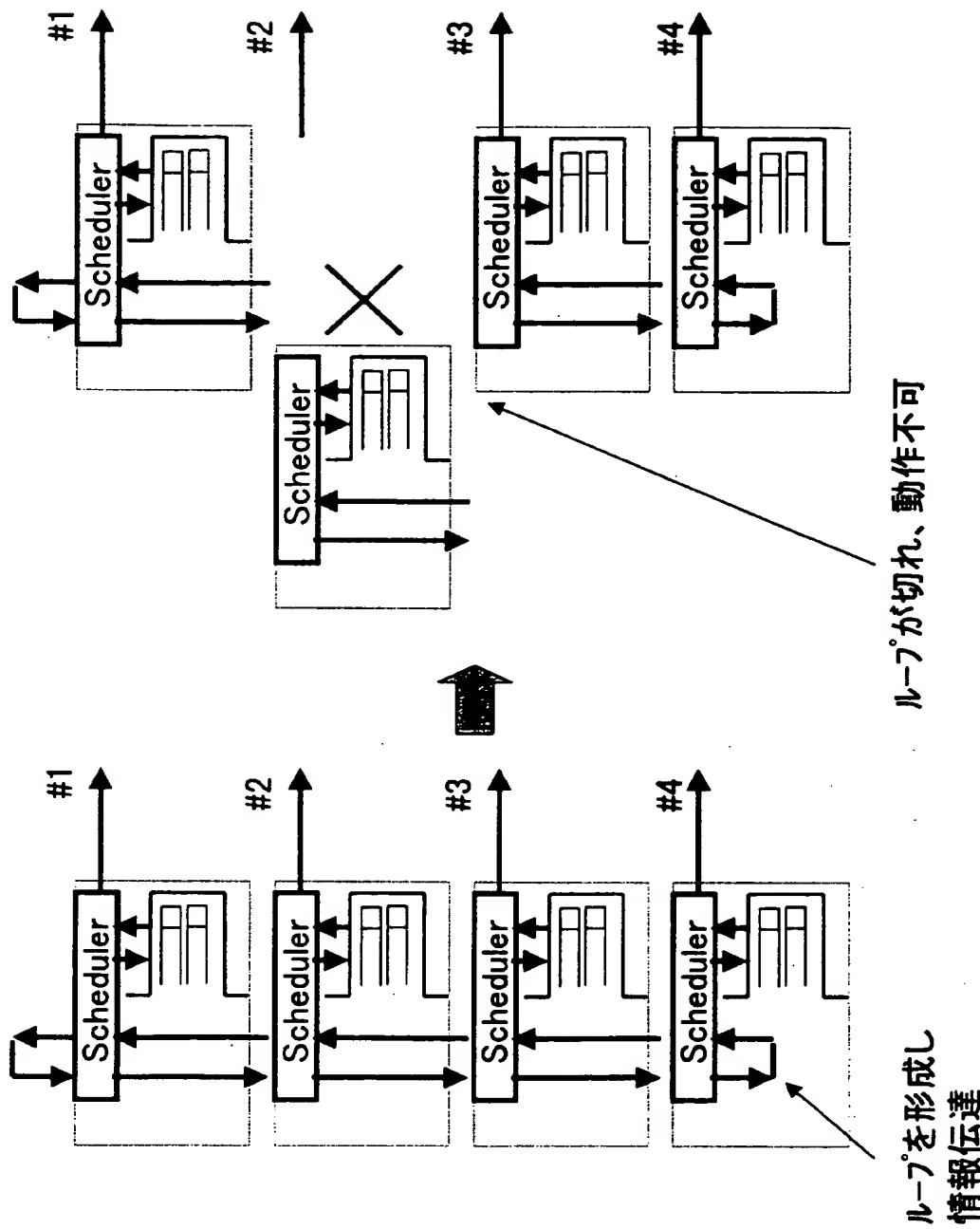


【図10】



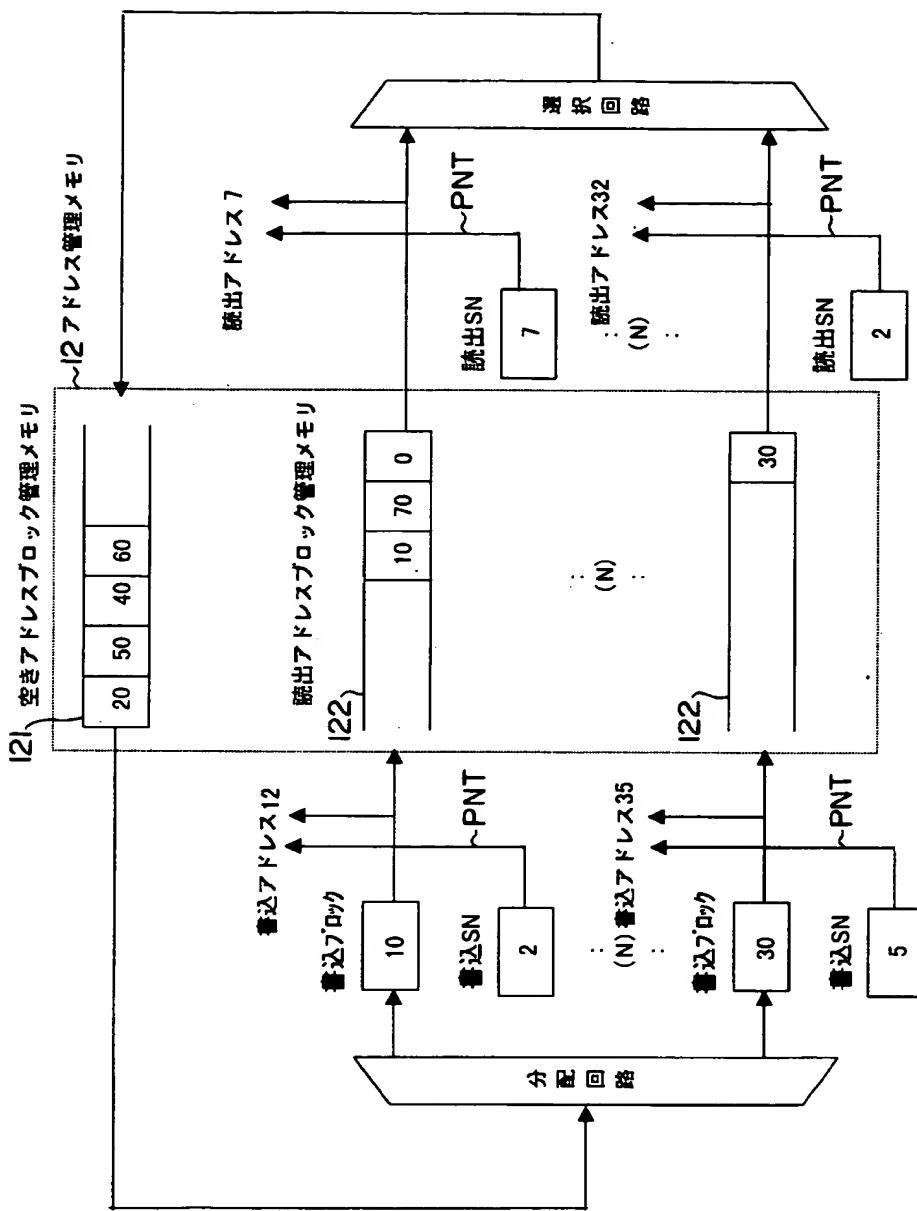
【図 11】

パケットスイッチ装置のスケジューリング情報の切断を説明するための図



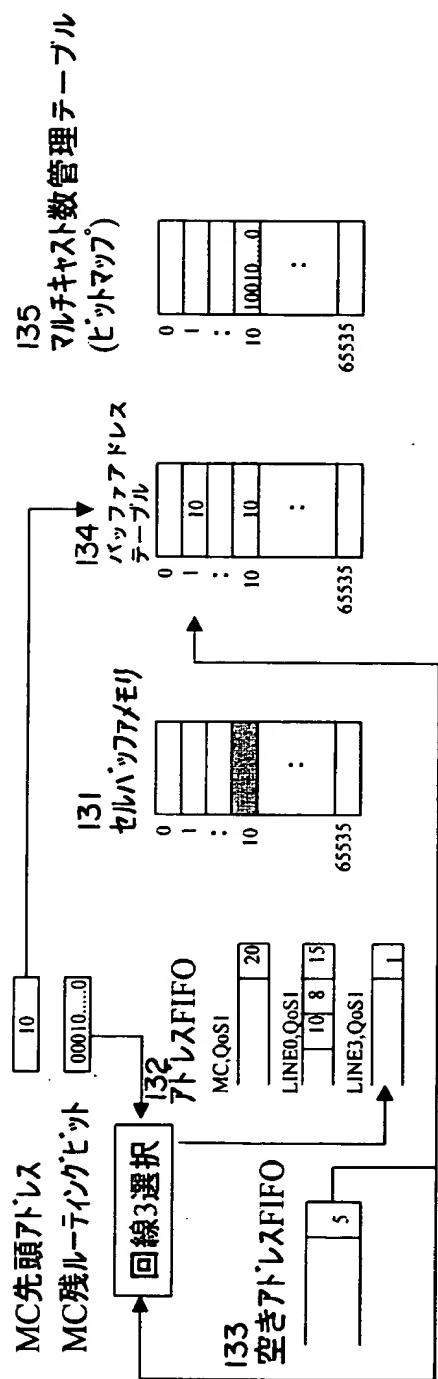
【図12】

本発明のパケットストリッピング装置のアドレスロック管理手法を説明するための図



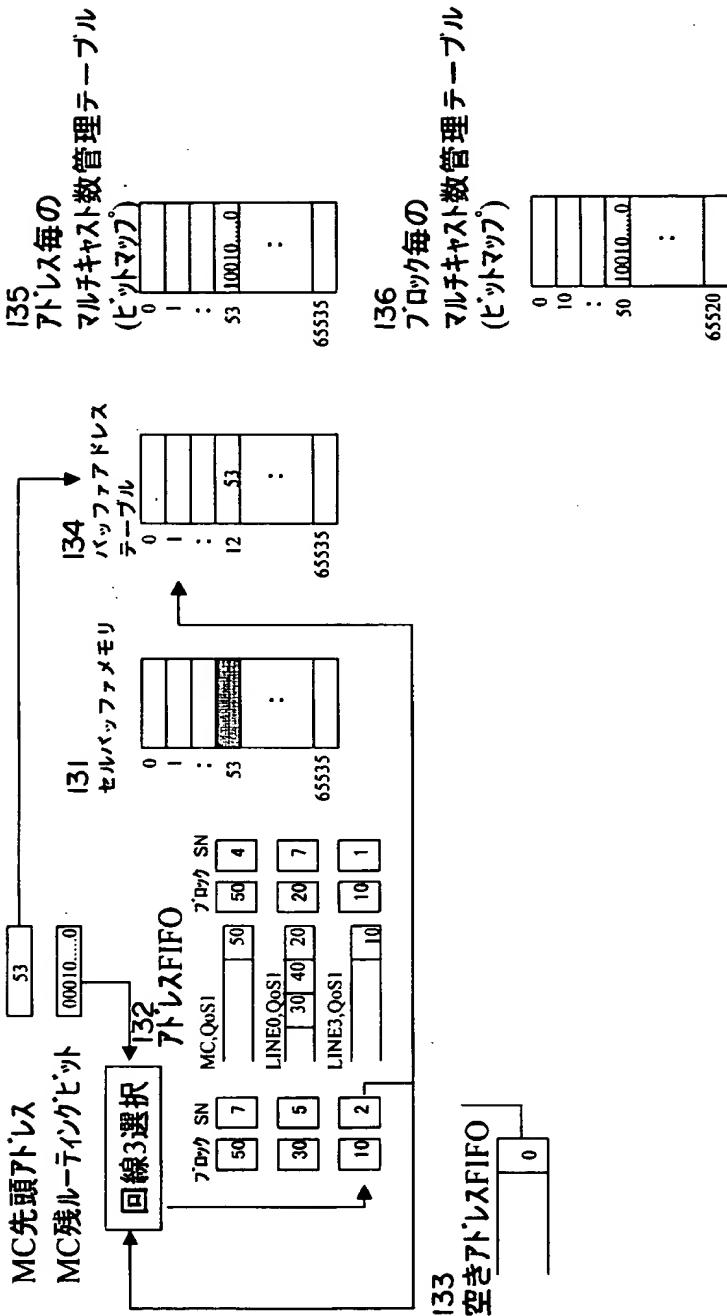
【図13】

アドレスリンクにおけるマルチキャストを説明するための図



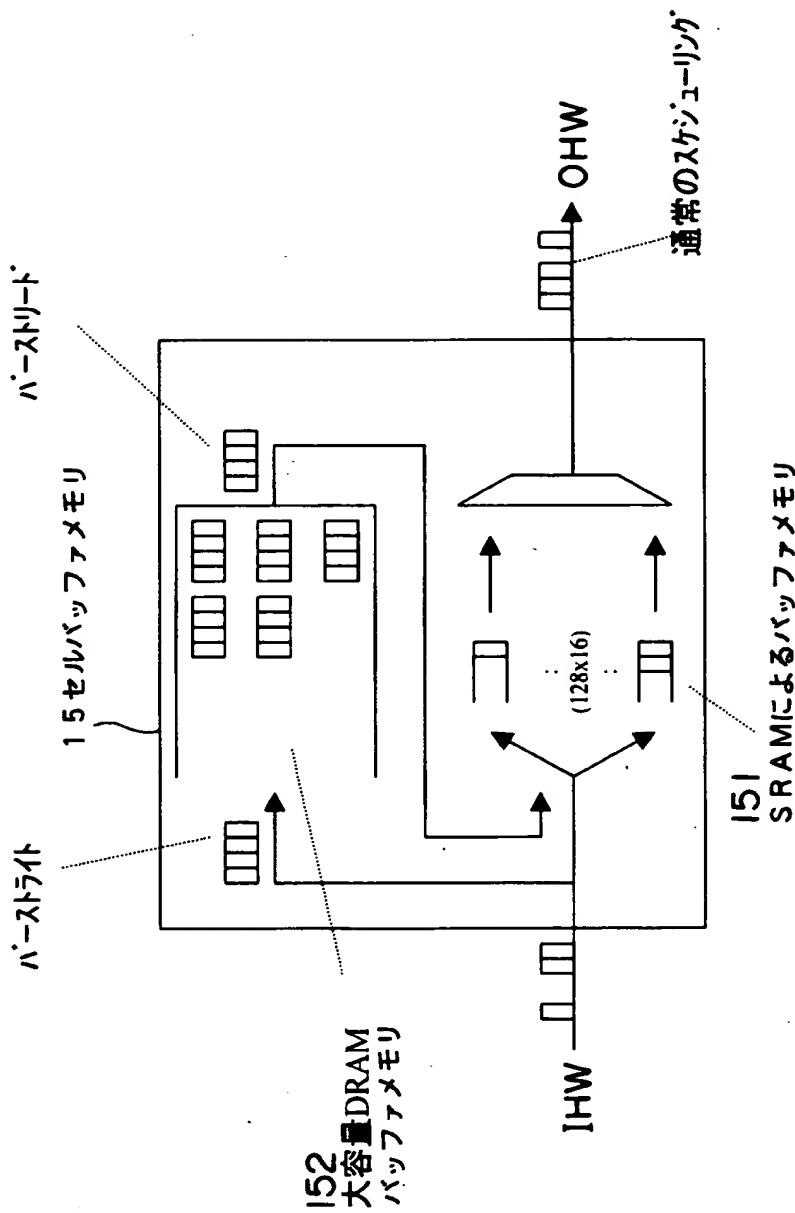
【図14】

ブロック管理手法におけるマルチキャストを説明するための図



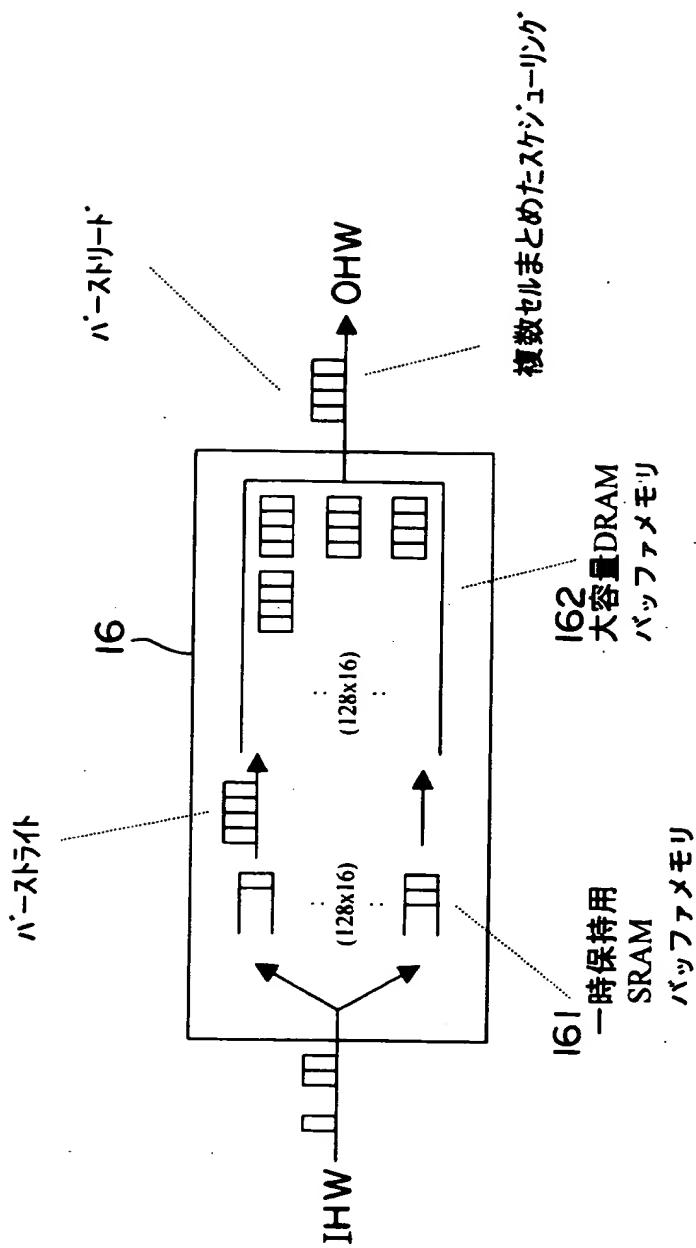
【図15】

セルバッファメモリの構成例を示すブロック図



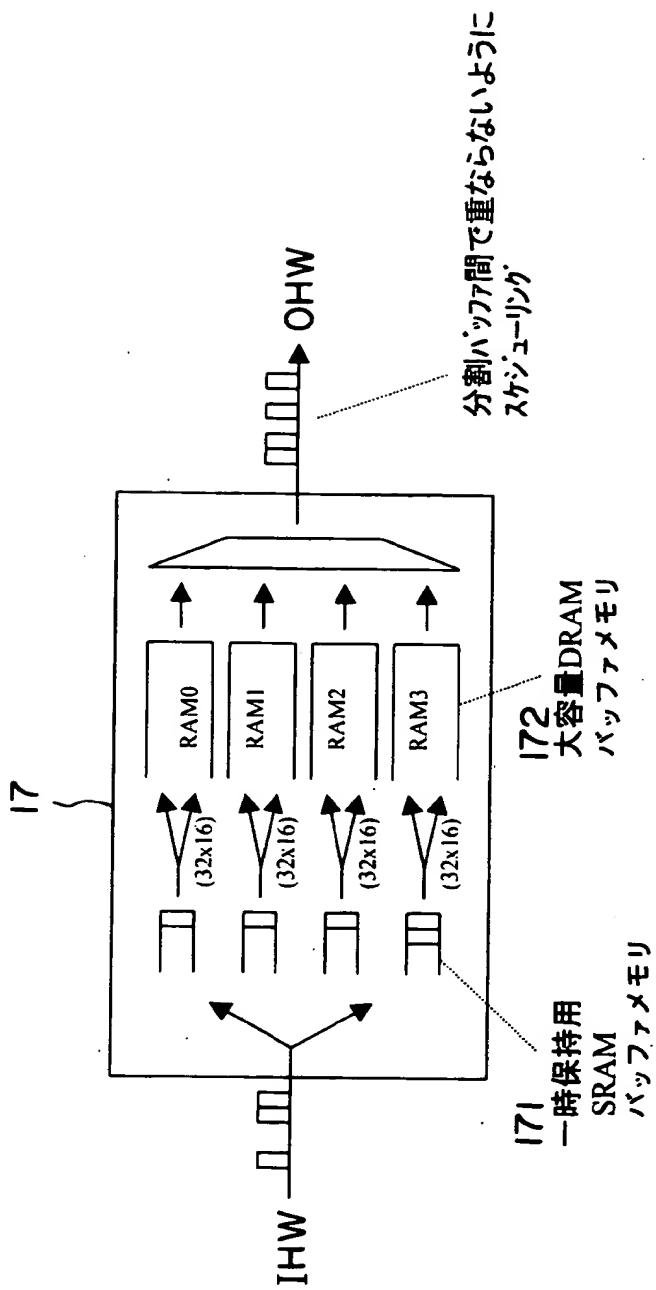
【図16】

セルバッファメモリの構成例を示すブロック図



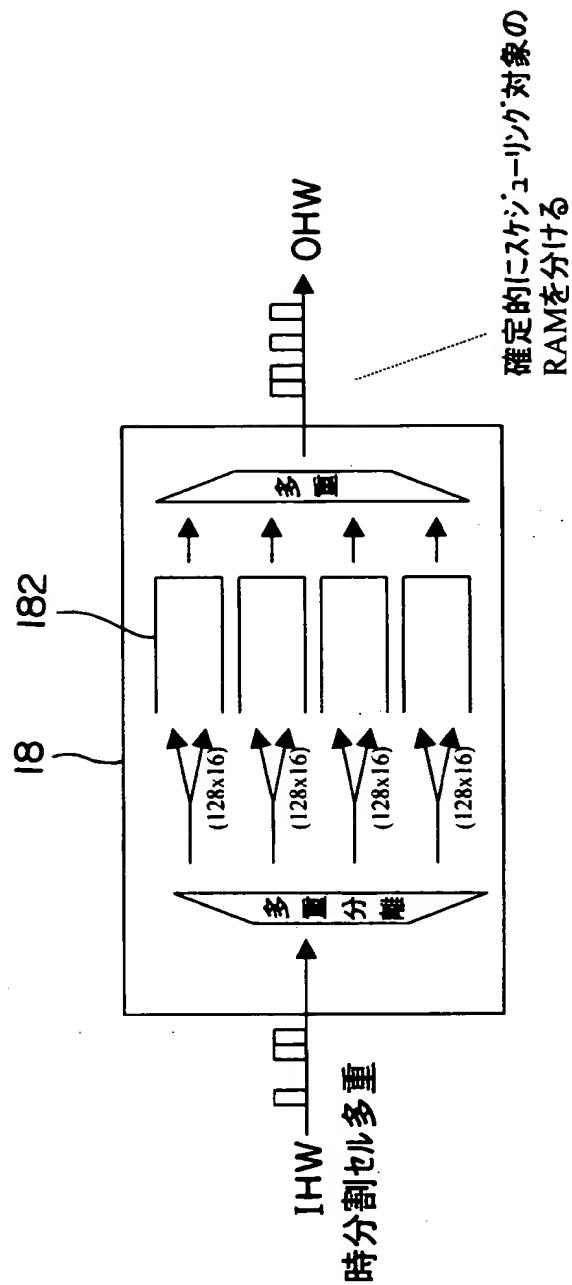
【図17】

セルバッファメモリの構成例を示すブロック図



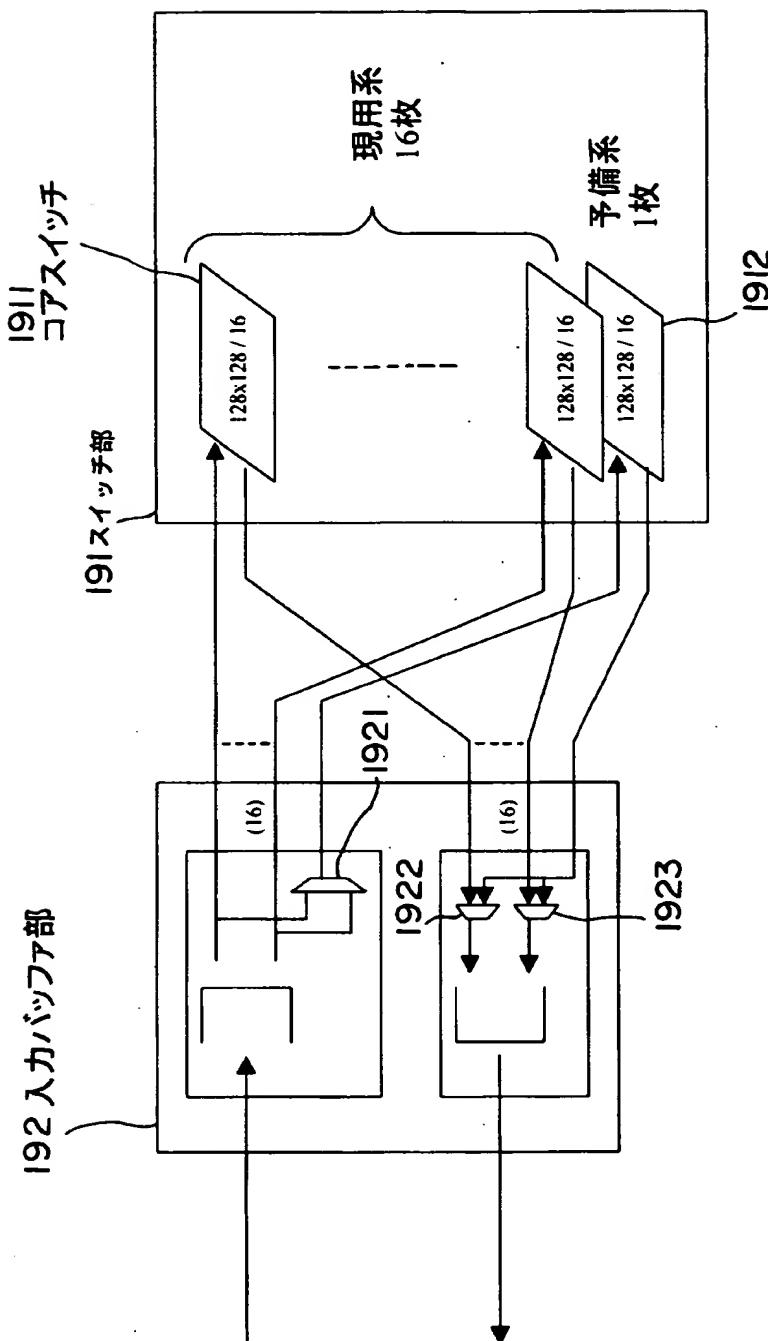
【図 18】

セルバッファメモリの構成例を示すブロック図



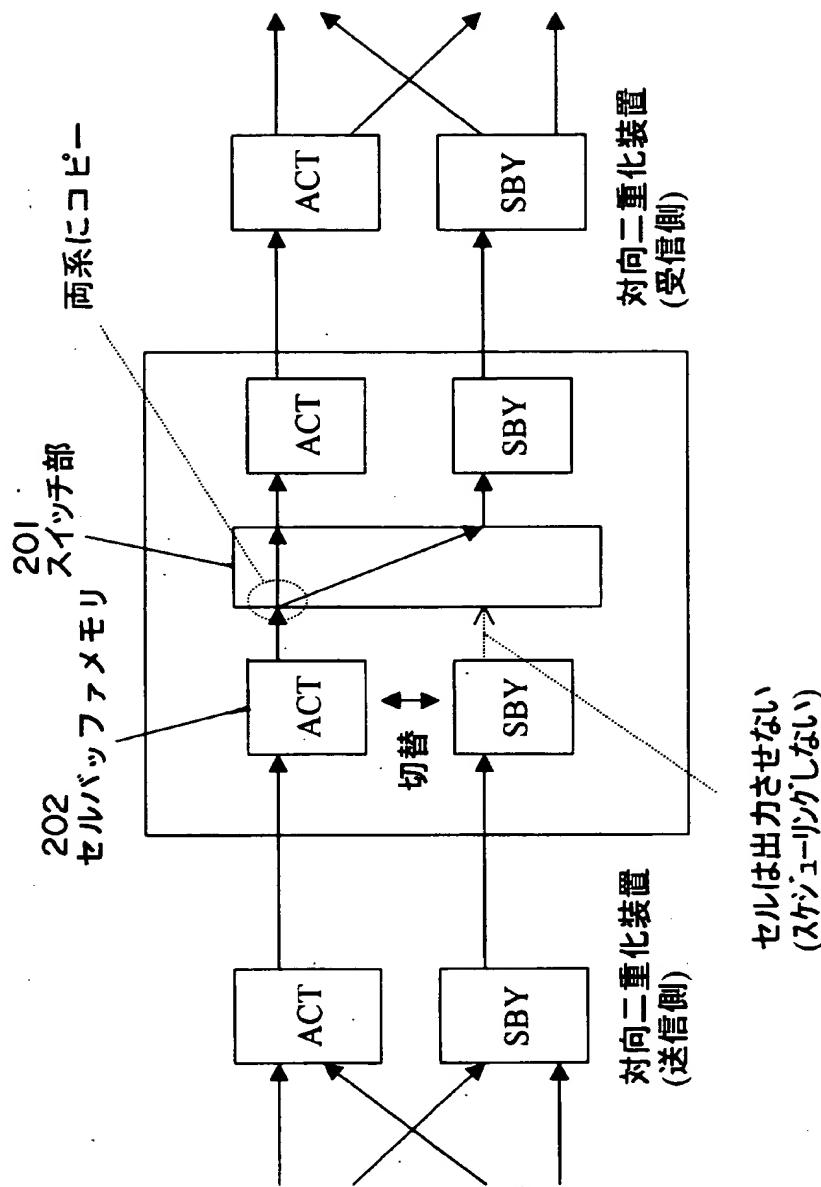
【図19】

コアスイッチの冗長構成例を示すブロック図



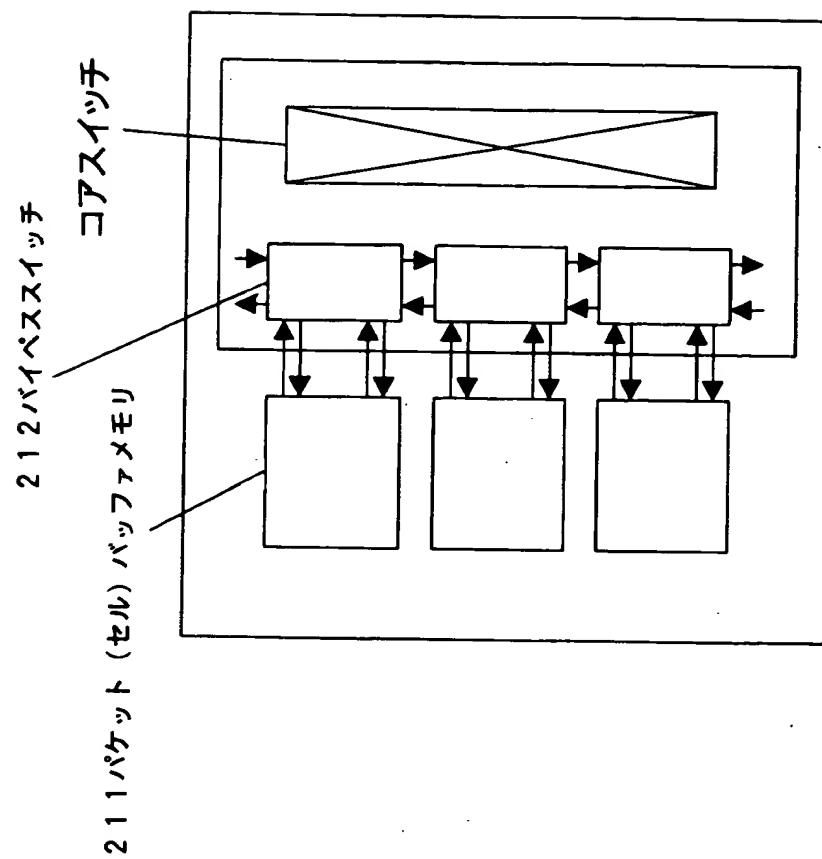
【図20】

セルバッファメモリの冗長構成例を示すブロック図

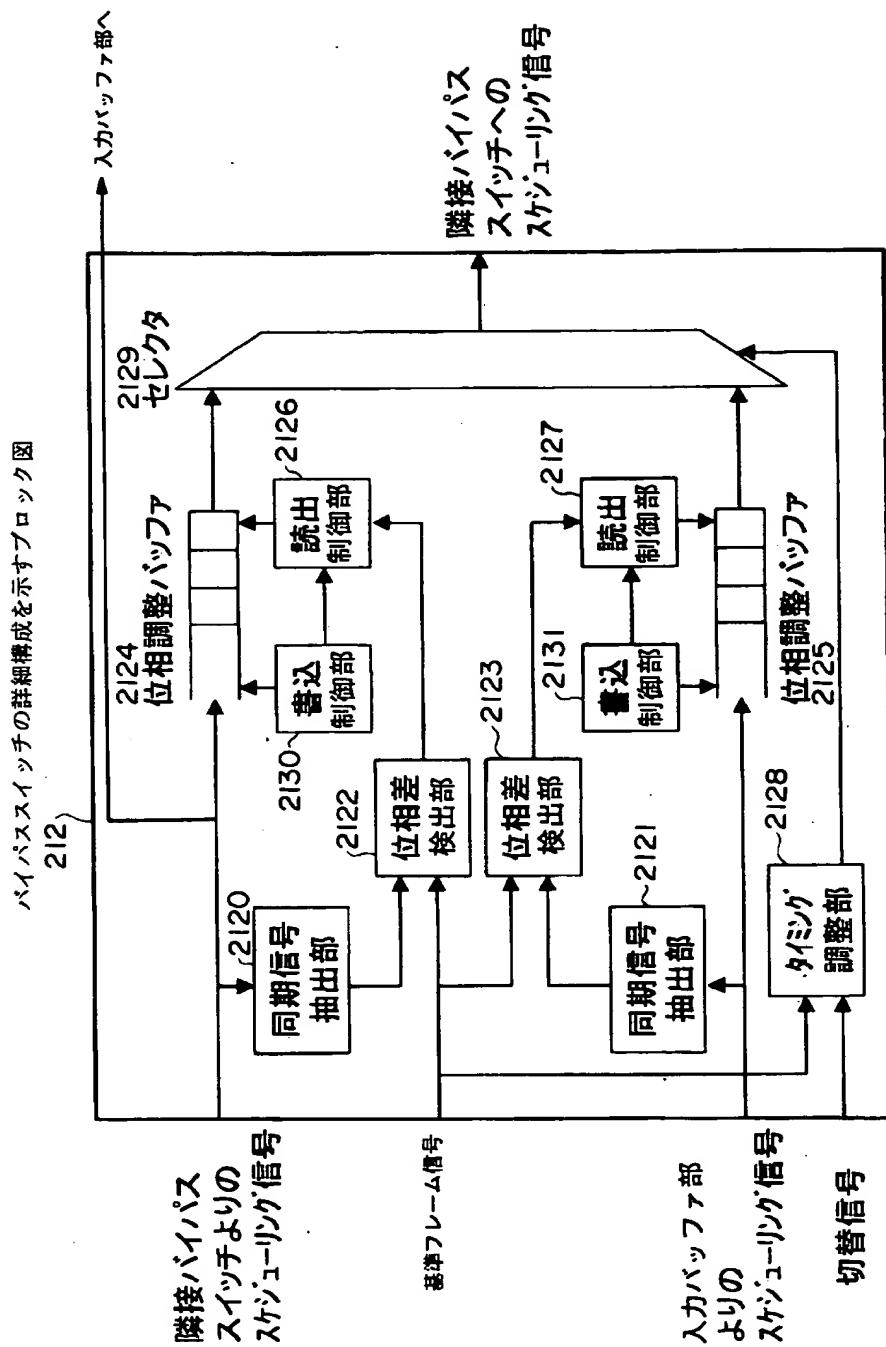


【図21】

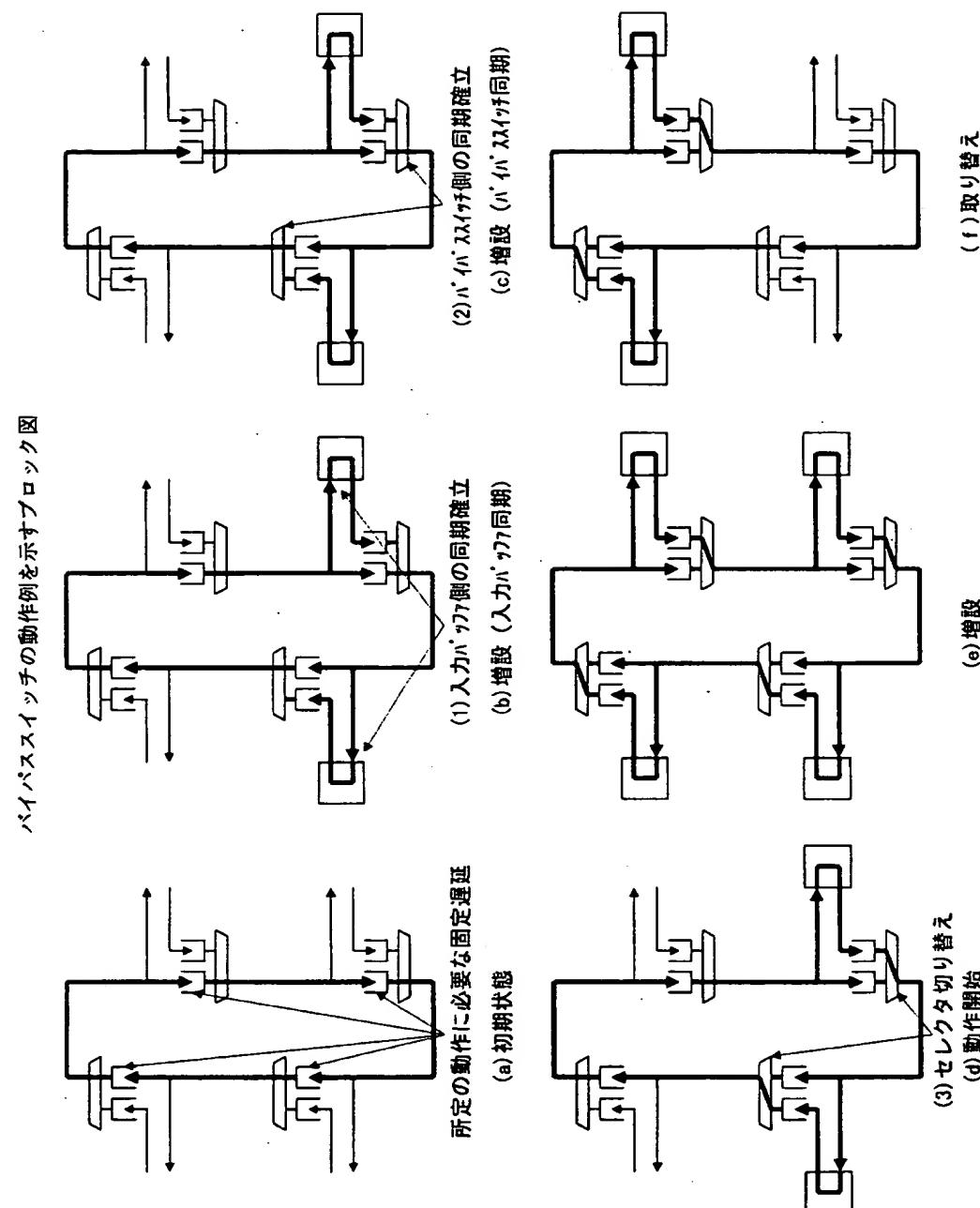
スケジューリング情報の切替構成を示すブロック図



【図22】

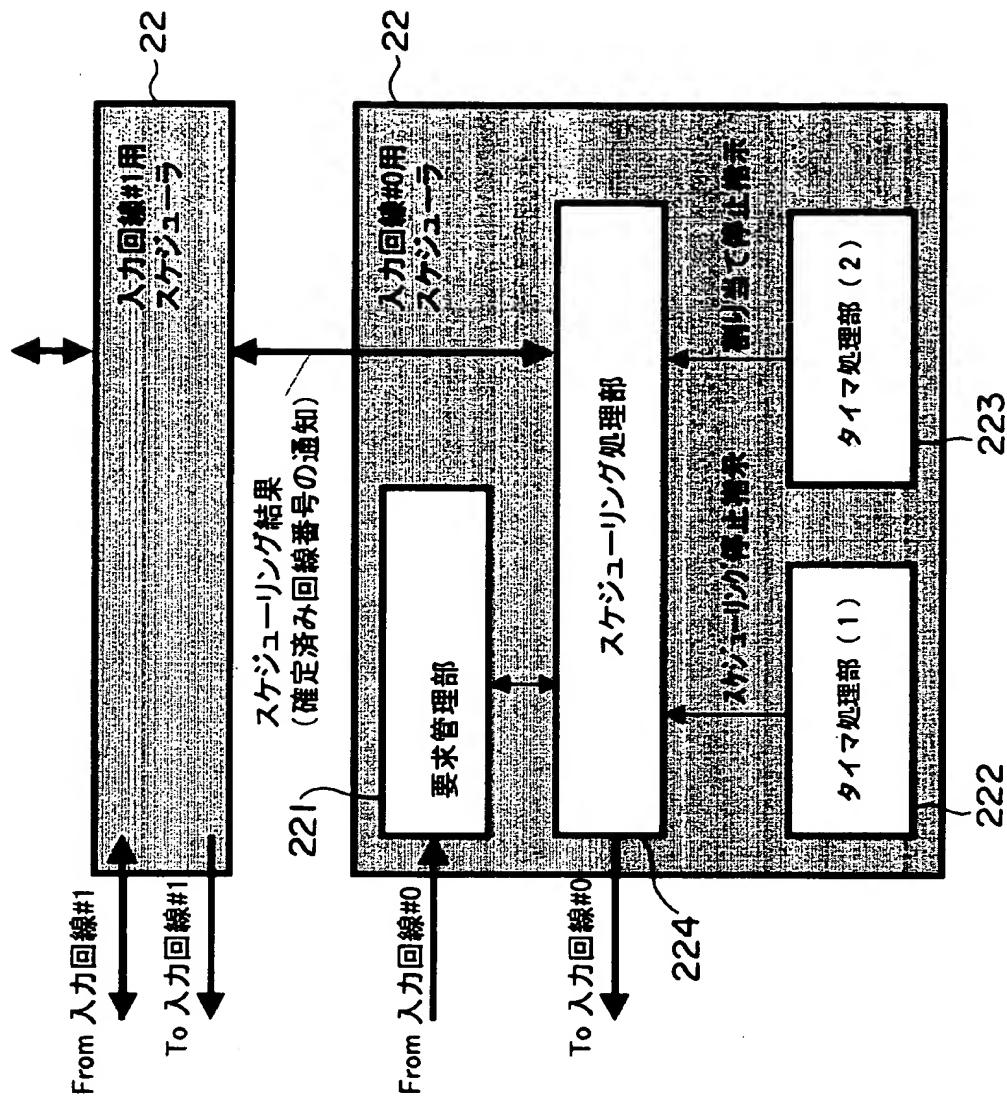


【図23】



【図24】

第1のスケジューラの構成を示すブロック図



【図 25】

書込動作を説明するための図

```

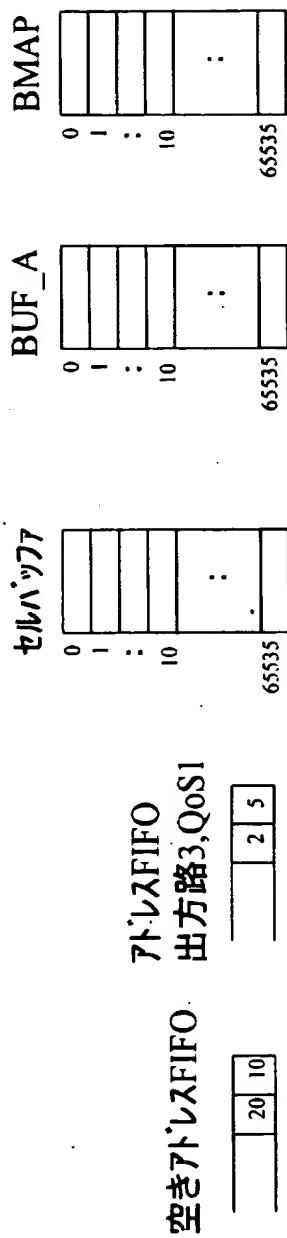
-- アドレス獲得
if INI_CNT < MAX -- アドレス生成カウントより獲得
    INI_CNT ++
    W_ADR = INT_CNT
else
    -- 空きアドレス FIFOより獲得
    W_ADR = EMP_S_PNT
    EMP_S_PNT = LINK(EMP_S_PNT)
endif
-- ポイント・リンク更新
if CNT_L(BUF) = 0 -- ベルヒッファが空の時の処理
    S_PNT(BUF) = W_ADR
    E_PNT(BUF) = W_ADR
else
    -- ベルヒッファが空でない時の処理
    LINK(E_PNT(BUF)) = W_ADR
    E_PNT(BUF) = W_ADR
endif
-- カウンタ更新
CNT_L(BUF) ++
CNT_S(QOS) ++
-- ヒッファアドレス更新
BUF_A(W_ADR) <= W_ADR
-- ビットマップ更新
if CELL(M) = 0 -- ユニキャストセルの時
    BMAP(W_ADR) <= BITMAP(CELL(UC-TAG))
else
    BMAP(W_ADR) <= CELL(MC-TAG)
endif

```

INI_CNT: 初期アドレス生成カウント
MAX: 使用バッファ量
W_ADR: 書き込みアドレス
EMP_S_PNT: 空きアドレス FIFO START ポインタ
LINK(x): アドレスにリンクされるアドレス
CNT_L(x): バッファxの個別バッファキューリング
CNT_M(x): QoS クラスx のマルチキャストバッファキューリング
CNT_S(x): QoS クラスx 共通バッファキューリング
S_PNT(x): バッファx の START ポインタ
E_PNT(x): バッファx の END ポインタ
BUF_A(x): アドレスのバッファアドレス
CELL(x): 入力セルのベルヒッファの値
BMAP(x): アドレスのルーティングビット(ビットマップ)
BITMAP(x): コードxをビットマップに変換

【図 26】

ユニキャスト書込動作を説明するための図

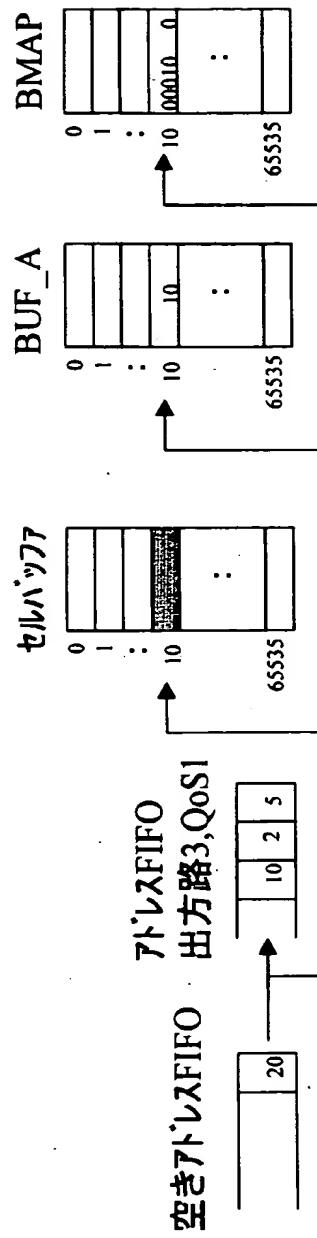


入力セル到着

- ユニキャストセル
- 出方路3
- QoS 1

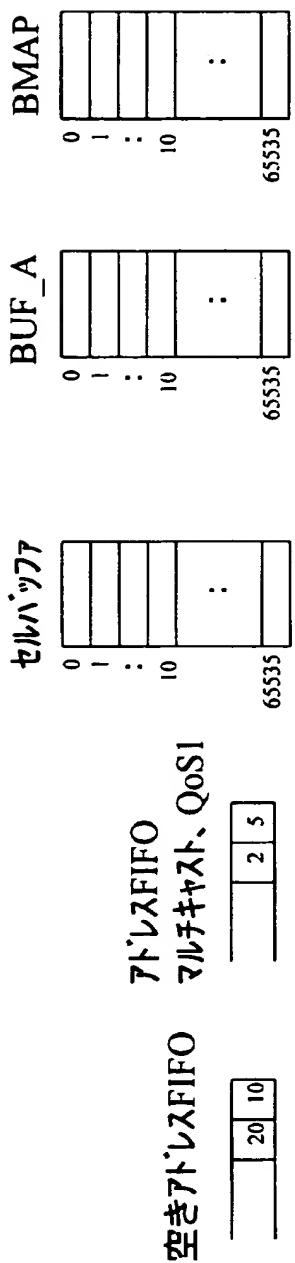


- 空きアドレス FIFO よりアドレス10を獲得
- セルバッファのアドレス10にセルを書き込み
- BUF_Aのアドレス10にアドレス10を設定
- BMAPのアドレス10のビット3をゼット
- アドレス FIFO の最後尾にアドレス10を追加

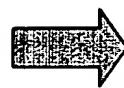


【図27】

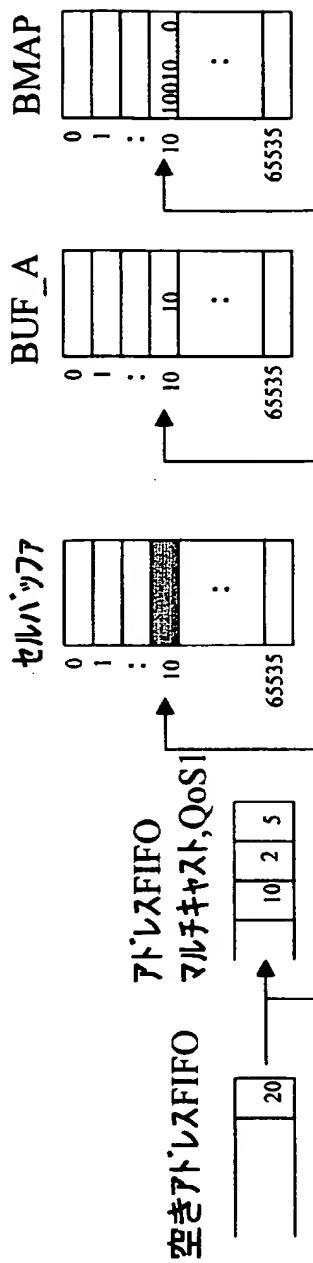
マルチキャスト書込動作を説明するための図



入力セル到着
 -マルチキャストセル
 -出方路0,3
 -QoS 1



- 空きアドレス FIFO よりアドレス10を獲得
- セルバッファのアドレス10にセルを書き込み
- BUF_A のアドレス10にアドレス10を設定
- BMAP のアドレス10のビット0,3をセット
- アドレス FIFO の最後尾にアドレス10を追加



【図 28】

マルチキャスト動作を説明するための図

```

--アドレス獲得          --マルチキャストセル目の場合
if MC_TOP_E = 0
--MC履歴更新
MC_TOP = S_PNT(MC_QOS)
MC_TOP_E = 1
MC_ADD = MC_TOP
MC_BMAP = BMAP(MC_TOP)

--ポイント更新
S_PNT(MC_QOS) = LINK(MC_TOP)
--カウント更新
CNT_M(MC_QOS) --
else
--マルチキャストセル目以降の場合
if INI_CNT < MAX --アドレス生成がカウントより獲得
INI_CNT ++
MC_ADD = INT_CNT
EMP_S_PNT: 空きアドレスFIFO STARTポインタ
LINK(x); アドレスにリンクされるアドレス
CNT_M(x); QoSクラスxのマルチキャストバッファキューラインク
CNT_S(x); QoSクラスx共通バッファキューラインク
INI_CNT: 初期アドレス生成カウント
MAX: 使用バッファ長
ADR_VAL: 空きアドレス有効
endif
endif
--カウント更新
CNT_S(MC_QOS) ++
endif

```

【図29】

マルチキャスト動作を説明するための図

```

-- ポイント更新
LINE = TOP(MC_BMAP)
BUF = LINE * 4 + MC_QOS
if CNT_L(BUF) = 0 -- セルバッファが空の時の処理
    S_PNT(BUF) = MC_ADD
    E_PNT(BUF) = MC_ADD
else -- セルバッファが空でない時の処理
    LINK(E_PNT(BUF)) = MC_ADD
    E_PNT(BUF) = MC_ADD
endif

-- カウンタ更新
CNT_L(BUF) ++
-- バッファアドレス保持
BUF_A(MC_ADD) = MC_TOP

-- ビットマップ更新
MC_BMAP -= BITMAP(LINE)
if MC_BMAP = 0 -- マルチキャスト終了判定
    MC_TOP_E = 0
endif

```

TOP(x): ビット列xのうち、0比特目からみて最初に
 が立っている比特番号を返す

LINE: 振り分け出方路番号

BUF: QoSクラスの個別バワキュー長

S_PNT(x): バッファのSTARTポインタ

E_PNT(x): バッファのENDポインタ

MC_ADD: マルチキャスト追加アドレス

LINK(x): アドレスにリンクされるアドレス

BUF_A(x): アドレスのバワアドレス

MC_TOP: マルチキャスト先頭アドレス

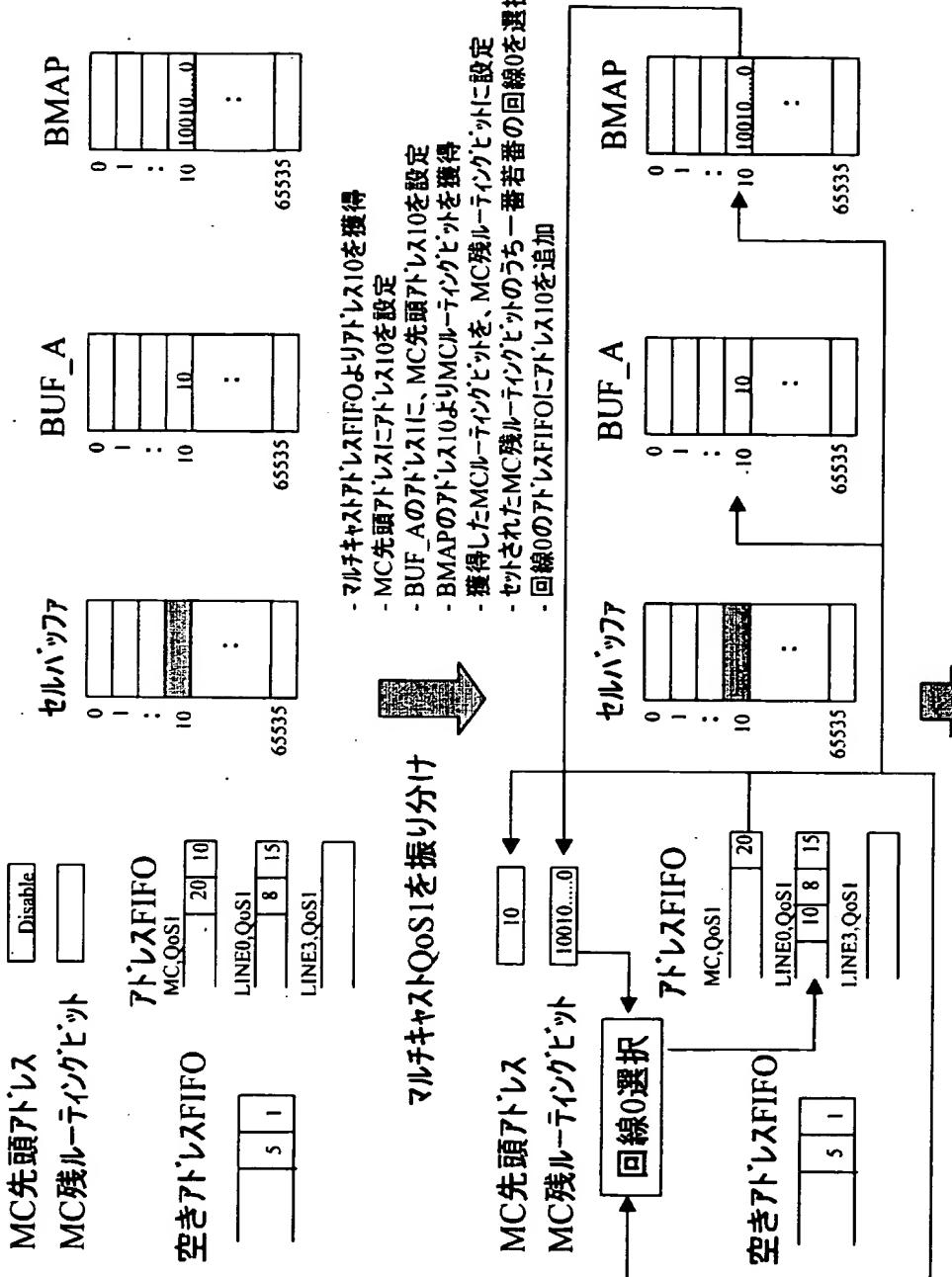
MC_BMAP: マルチキャスト残ルーティングビット(ビットマップ)

BITMAP(x): コードxをビットマップに変換

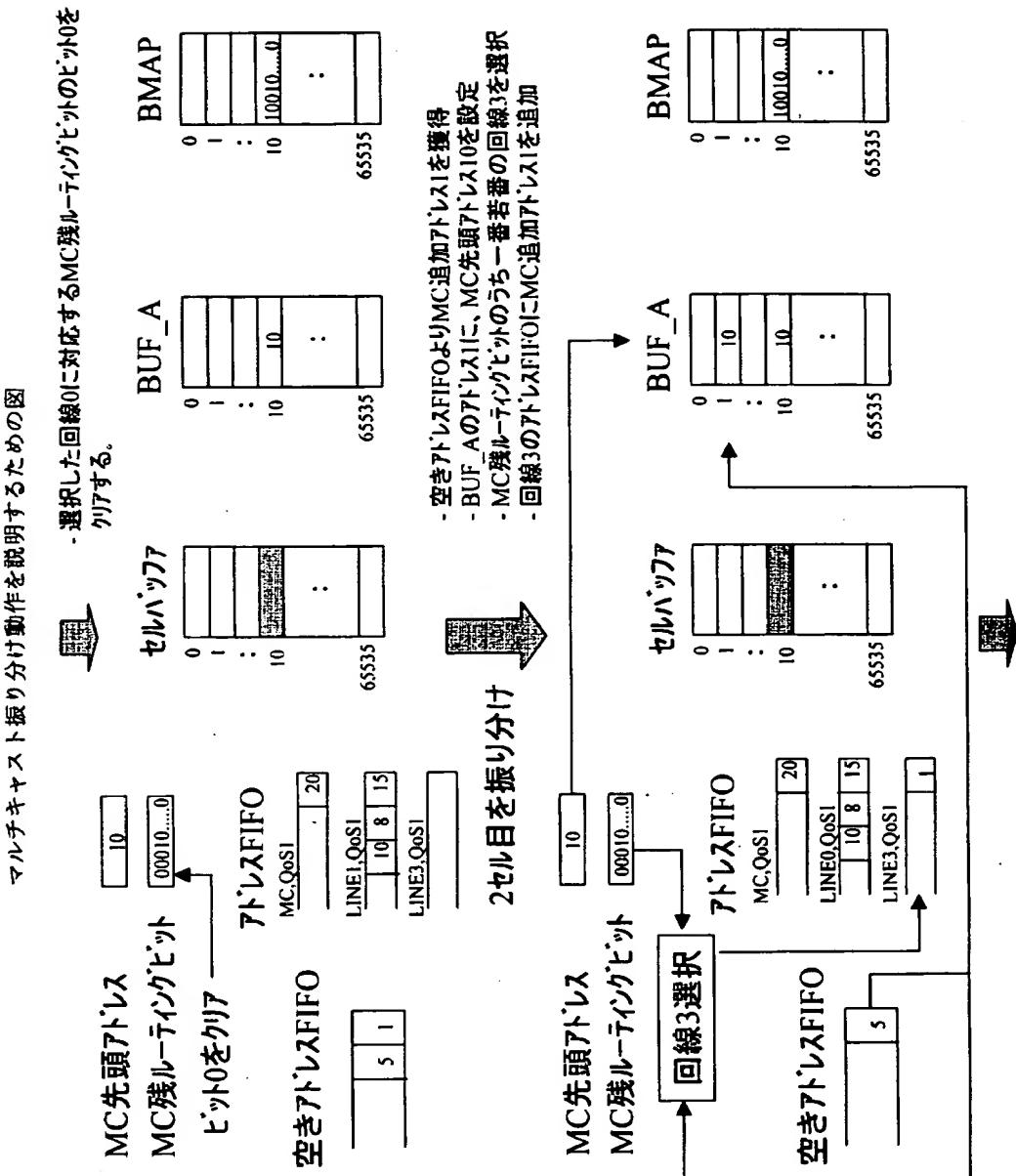
MC_TOP_E: マルチキャスト先頭アドレス有効

【図30】

マルチキャスト振り分け動作を説明するための図

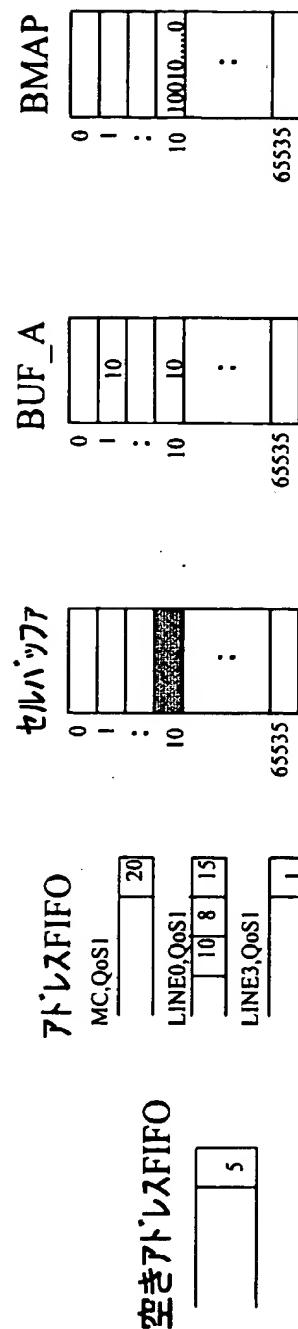
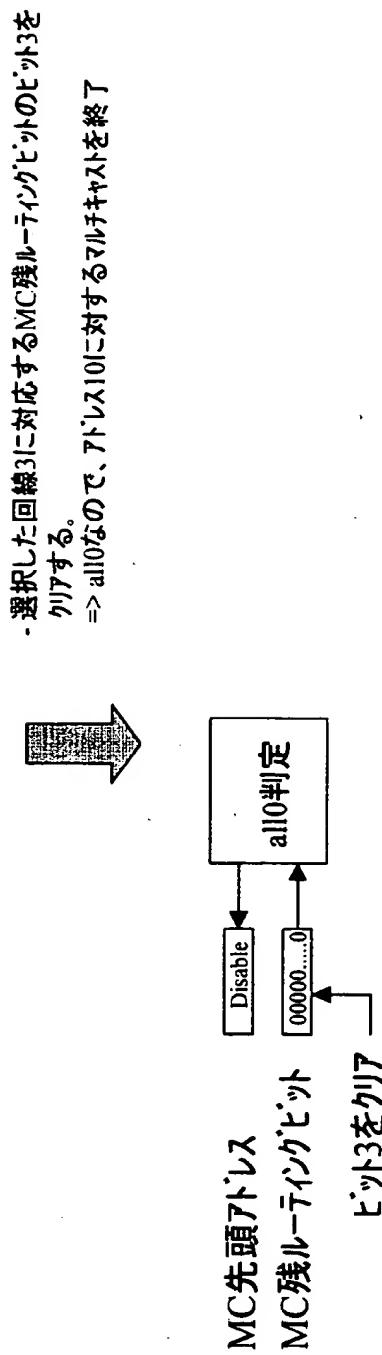


【図31】



【図3-2】

マルチキャスト振り分け動作を説明するための図



【図 3 3】

読み出動作を説明するための図

```
--アドレス獲得
BUF = LINE * 4 + QOS
R_ADR = S_PNT(BUF)
BUF_ADR = BUF_A(R_ADR)
--ビットマップ獲得
BMAP = BMAP(BUF_ADR)
```

```
--読み出しアドレス返却判定
if R_ADR <> BUF_ADR
    LINK(EMP_E_PNT) = R_ADR
    EMP_E_PNT = R_ADR
--カウンタ更新
    CNT_S(QOS) --
endif

--ポインタ更新
S_PNT(BUF) = LINK(S_PNT(BUF))

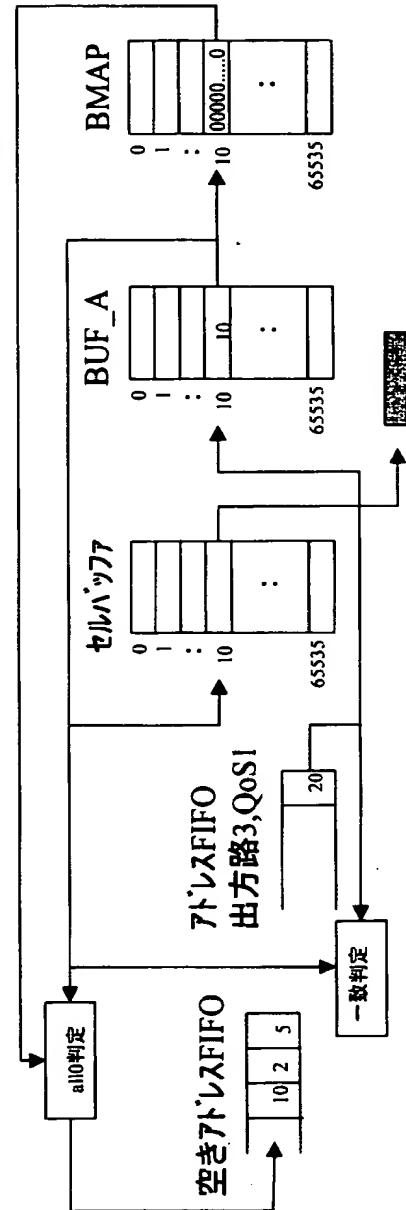
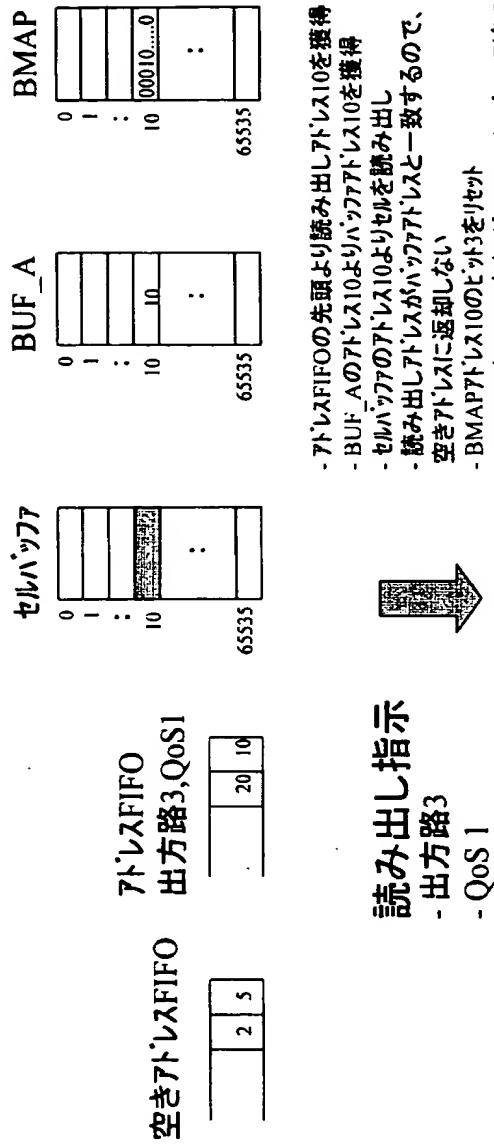
--ビットマップ更新
BMAP = BITMAP(LINE)
if BMAP = 0
--バッファアドレス返却
    LINK(EMP_E_PNT) = BUF_ADR
    EMP_E_PNT = BUF_ADR
--カウンタ更新
    CNT_S(QOS) --
endif

--カウンタ更新
CNT_L(BUF) --
```

読み出しアドレス
バッファxのSTARTポインタ
読み出しバッファ番号
読み出し方路番号
読み出しQoS番号
アドレスのバッファアドレス
アドレスのルーティングビット(ビットマップ)
バッファアドレスに対応するルーティングビット
(ビットマップ)
アドレスxにリンクされるアドレス
空きアドレス FIFO ENDポインタ
QoSカラスx共通バッファキューランク
QoSカラスxの個別バッファキューランク
BITMAP(x): コードxをビットマップに変換

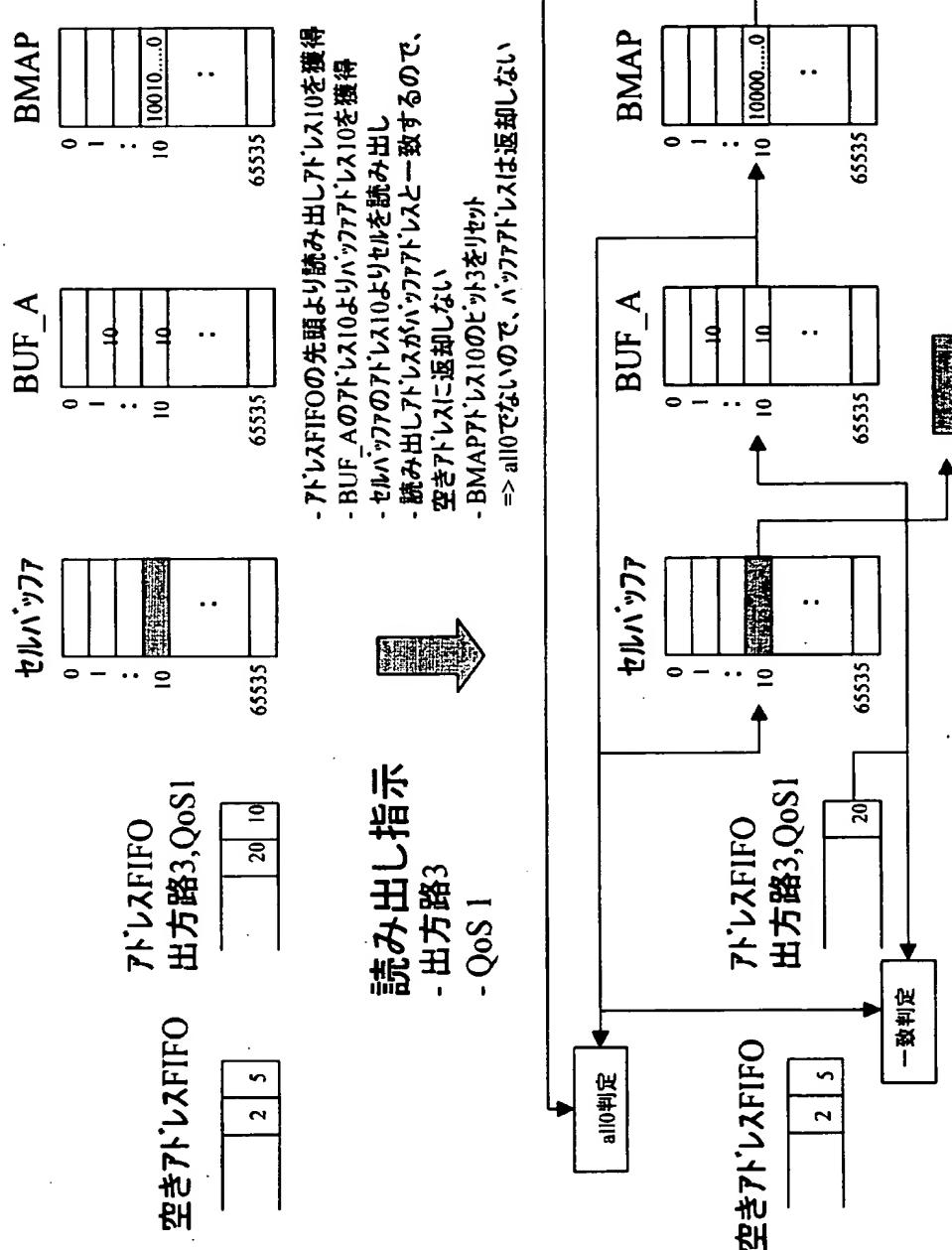
【図34】

読み出動作を説明するための図



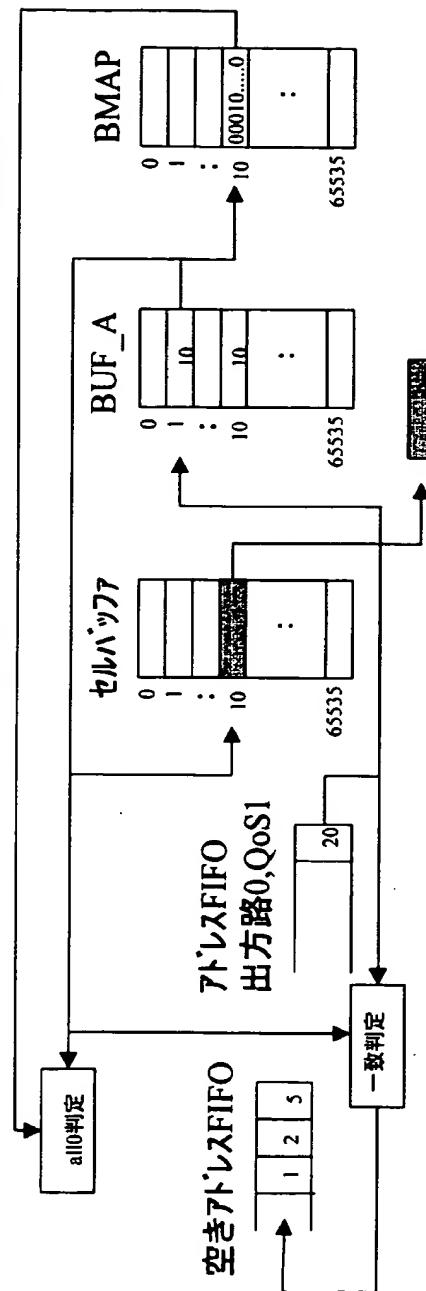
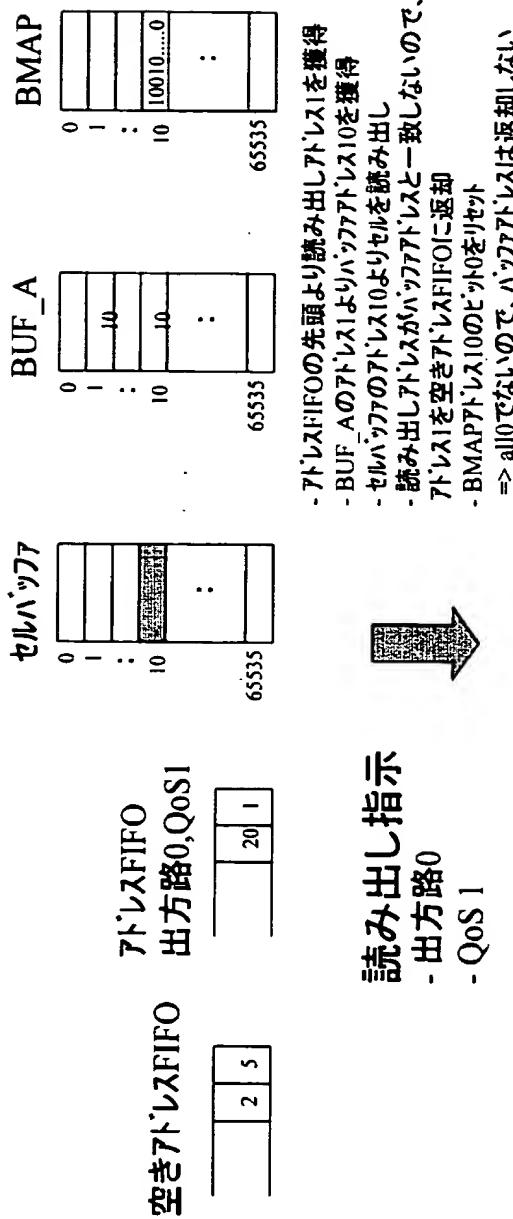
【図35】

読み出動作を説明するための図



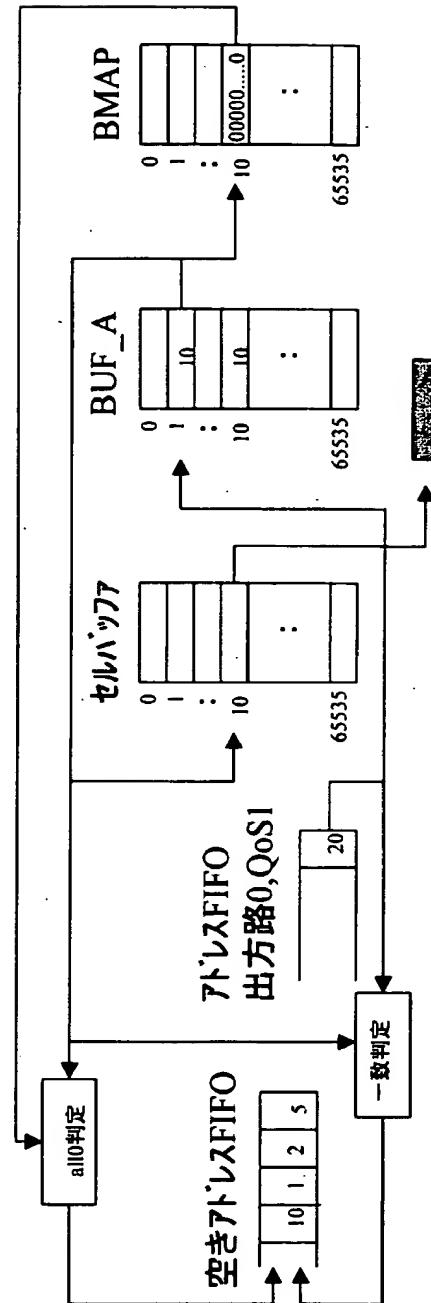
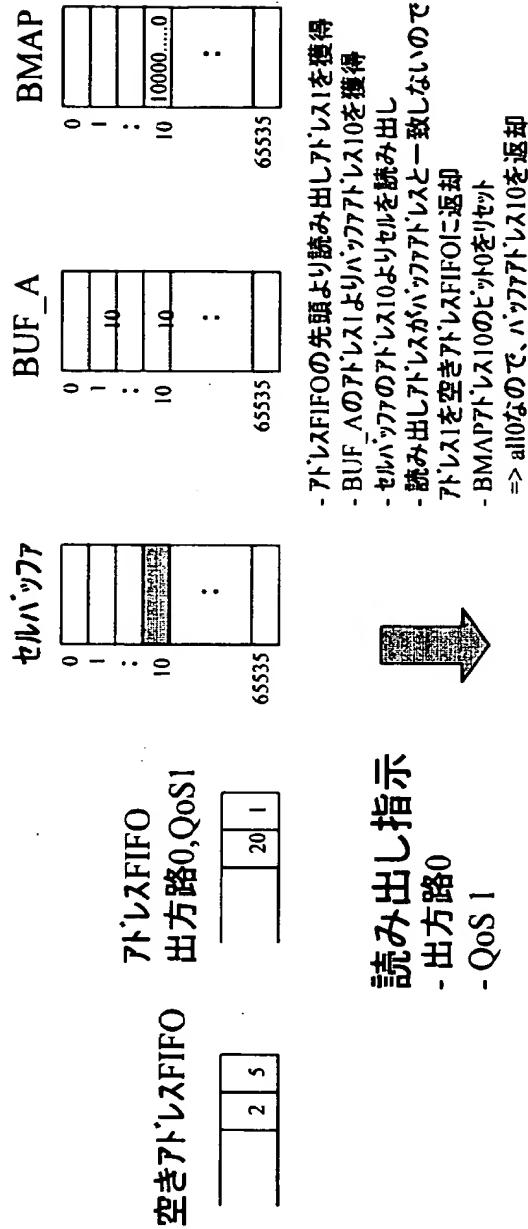
【図36】

読み出動作を説明するための図



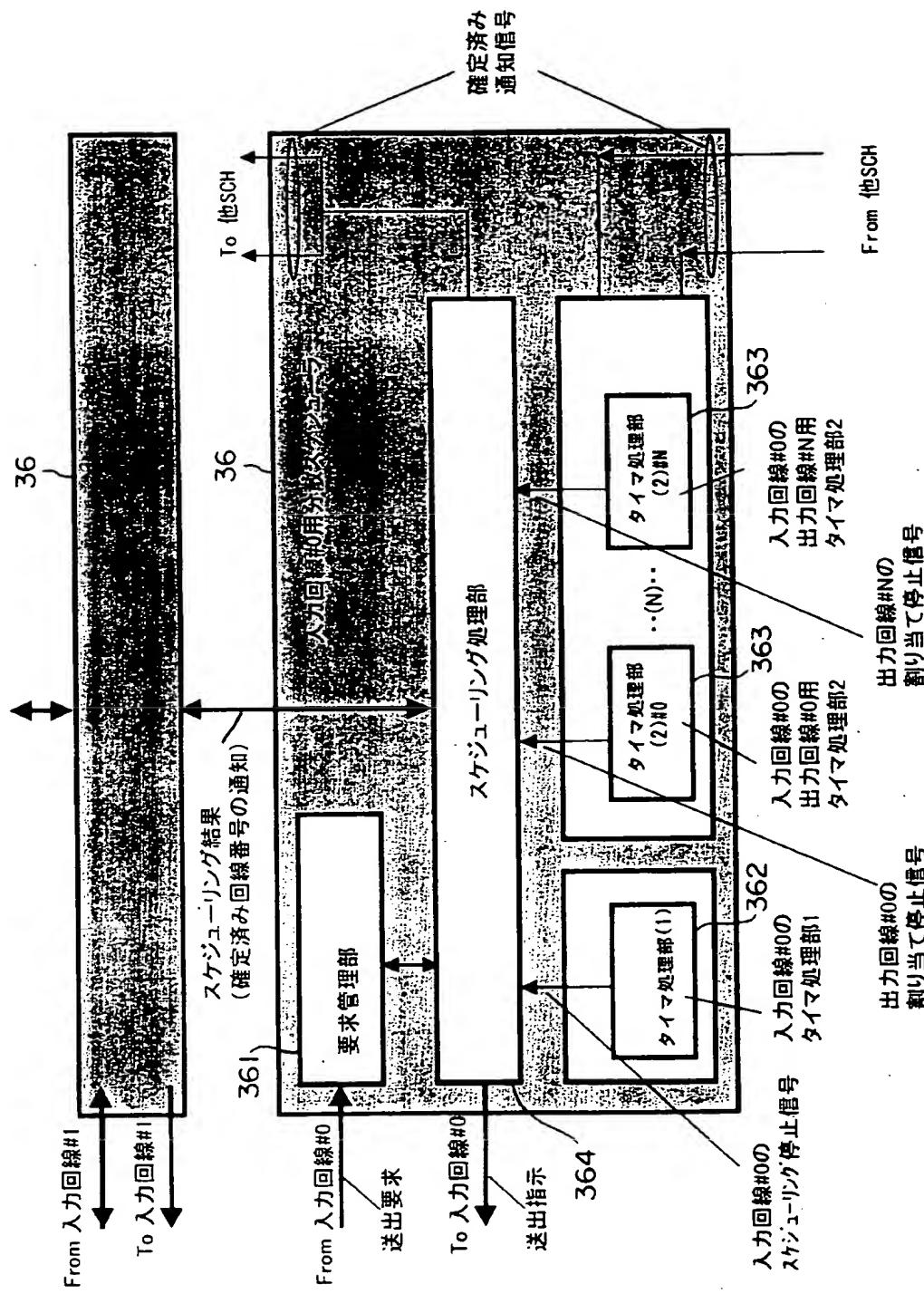
【図 37】

読み出動作を説明するための図



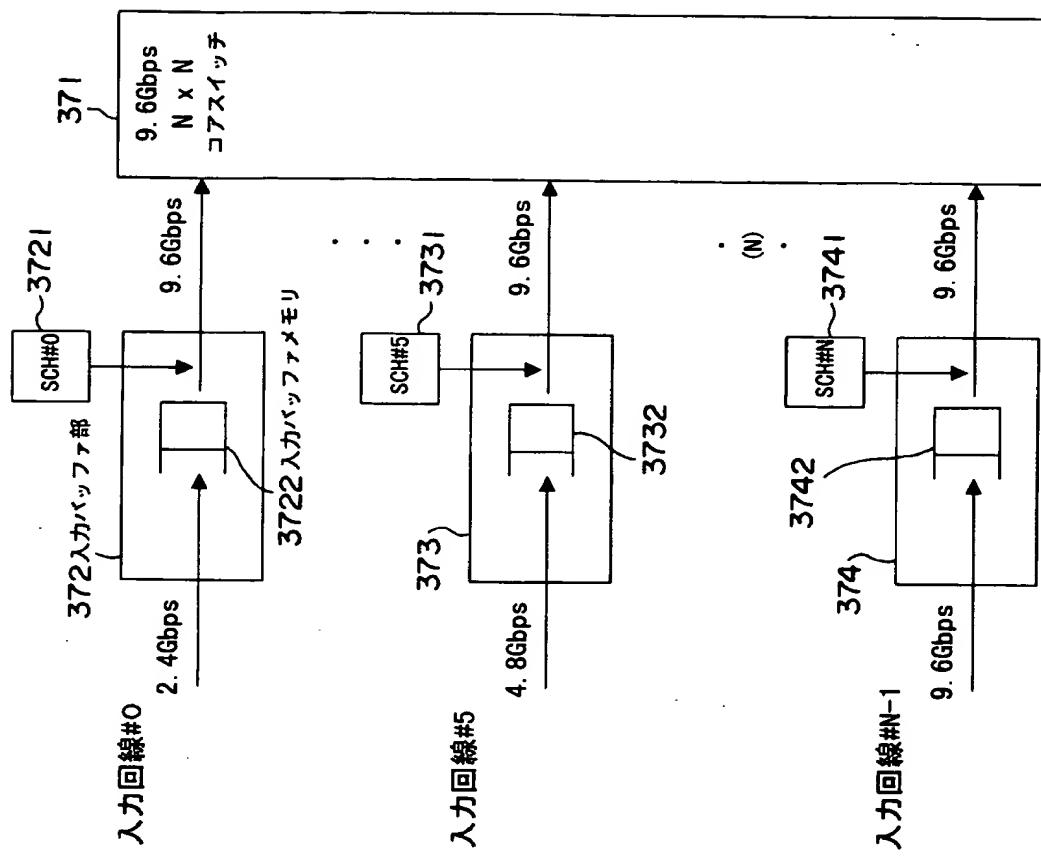
【図38】

第2のスケジューラの構成を示すブロック図

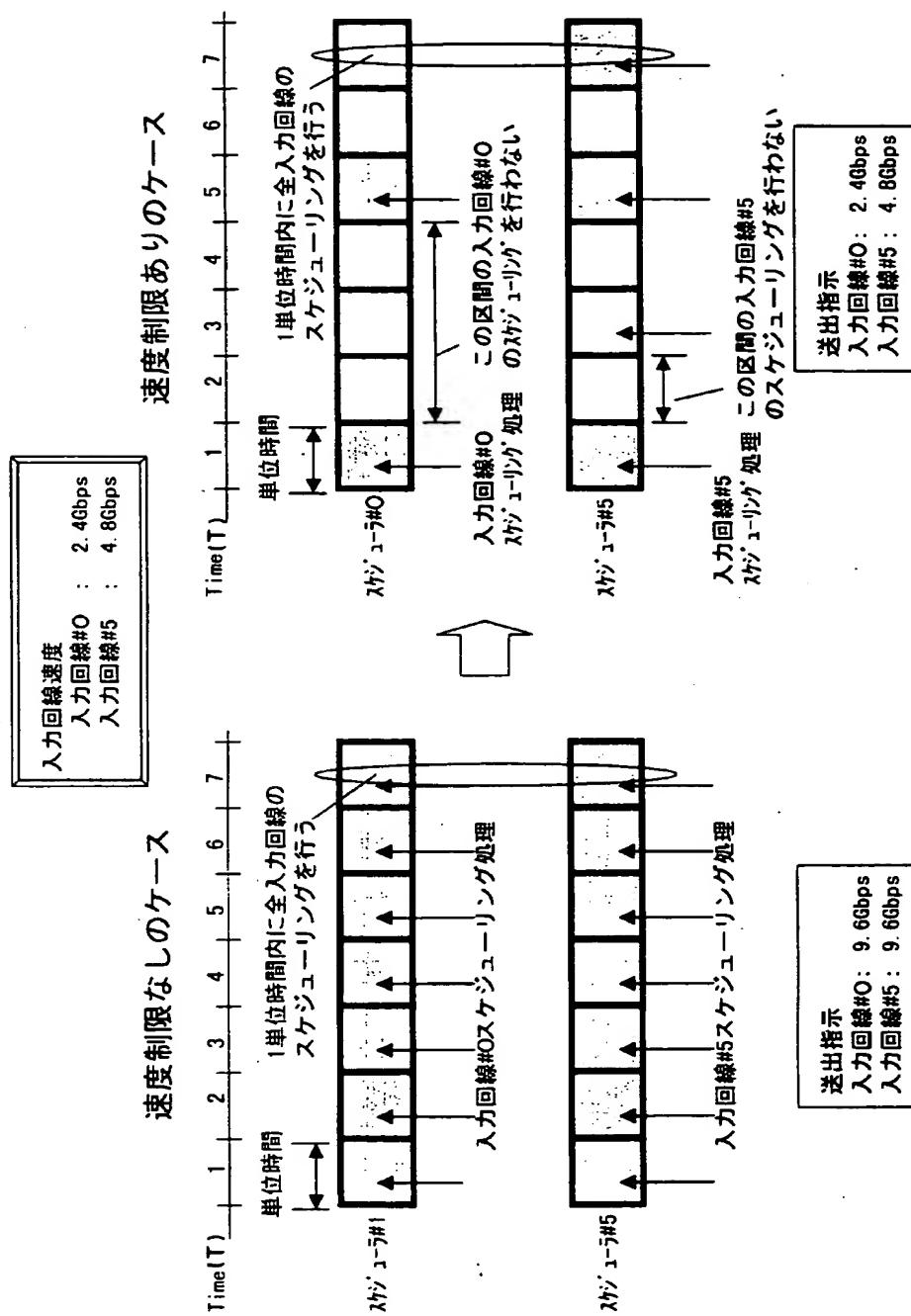


【図39】

入力バッファ部の処理能力の問題を説明するための図

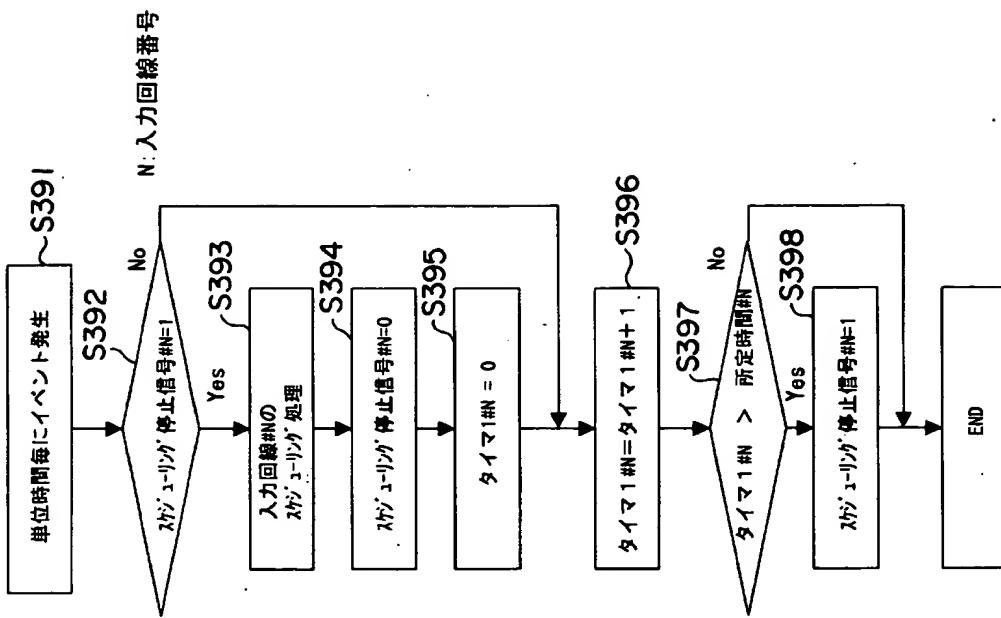


【図40】



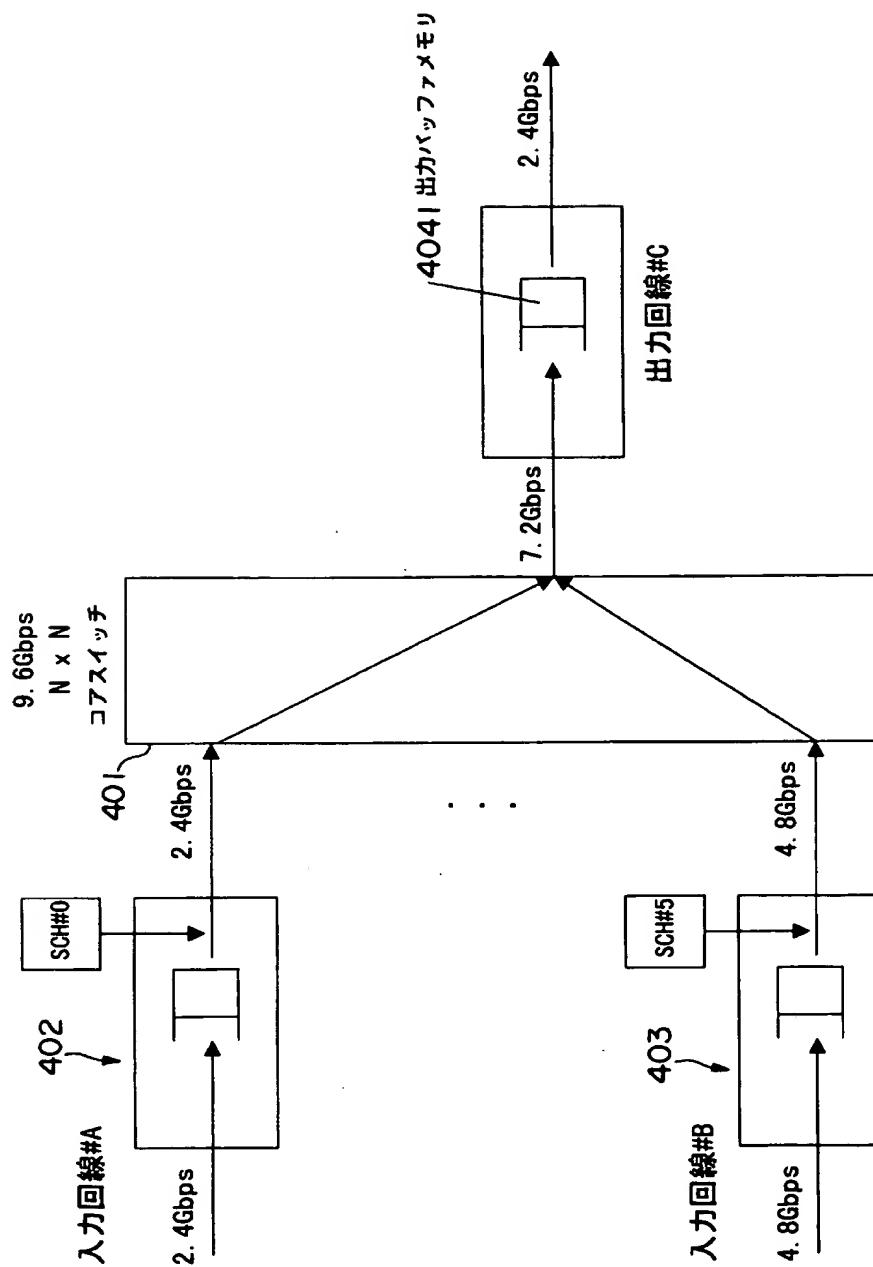
【図4-1】

入力回線速度に応じたスケジューリング処理手順を説明するための図

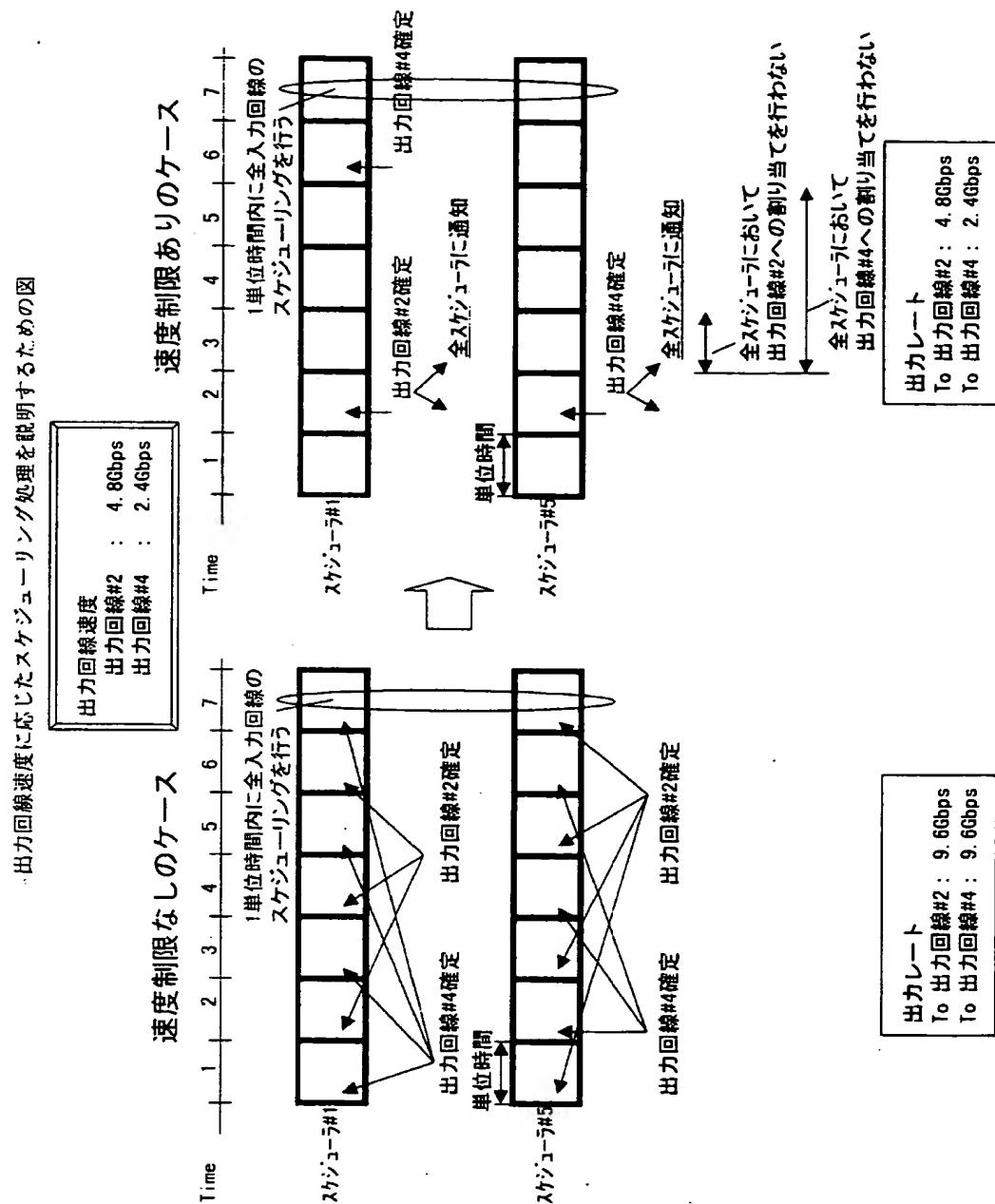


【図42】

出力トラヒック流の問題を説明するための図

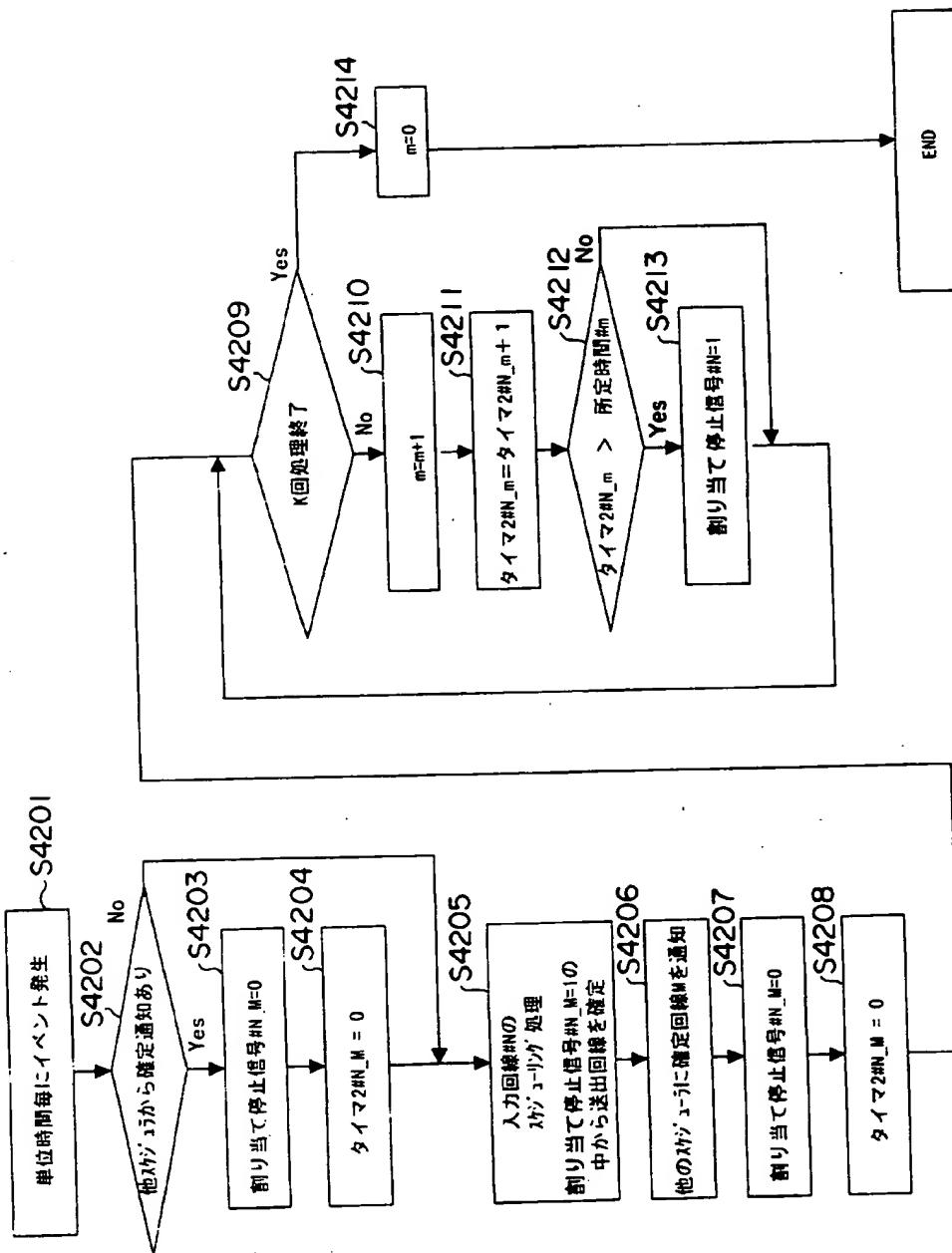


〔図43〕



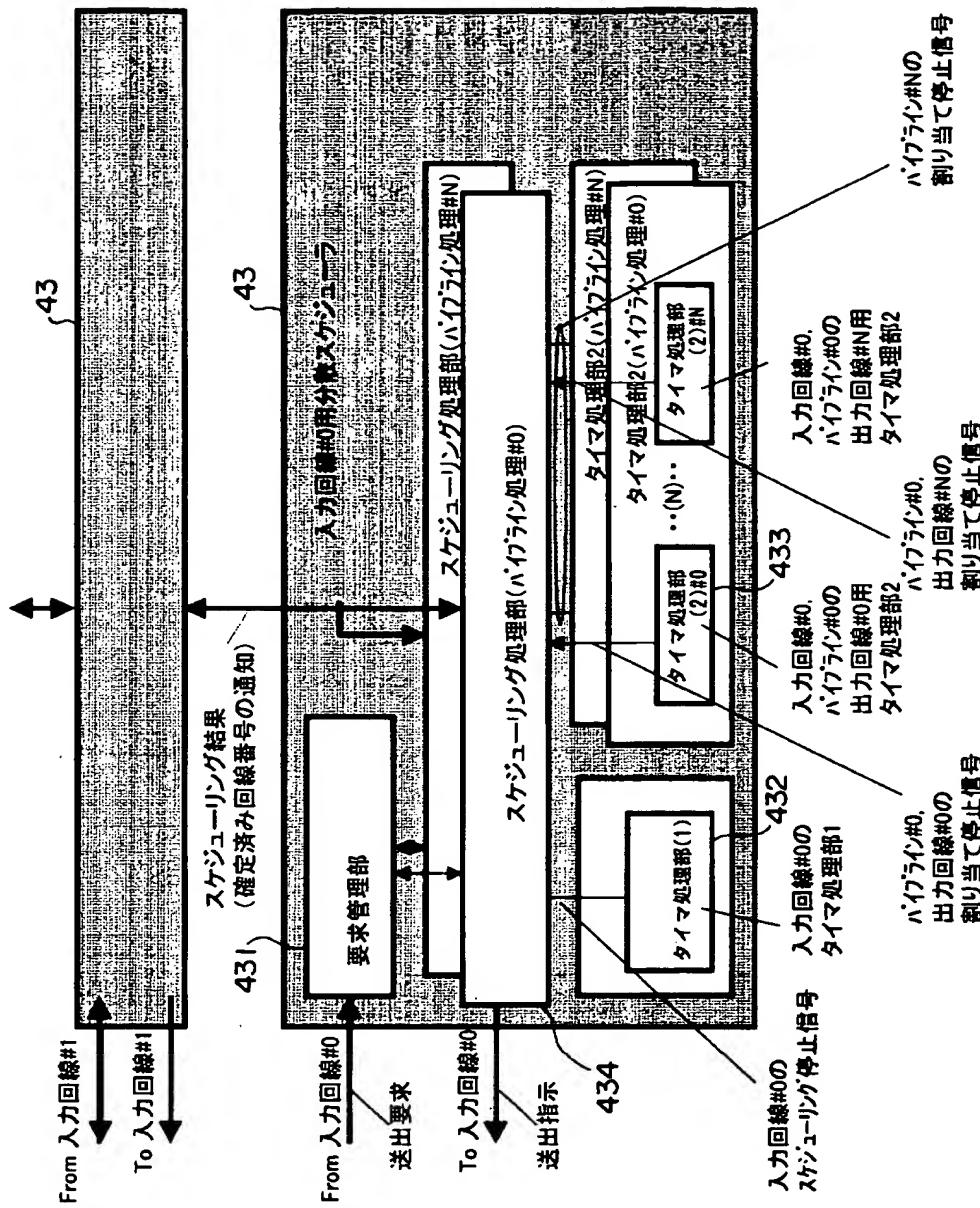
【図44】

出力回線速度に応じたスケジューリング処理手順を説明するための図



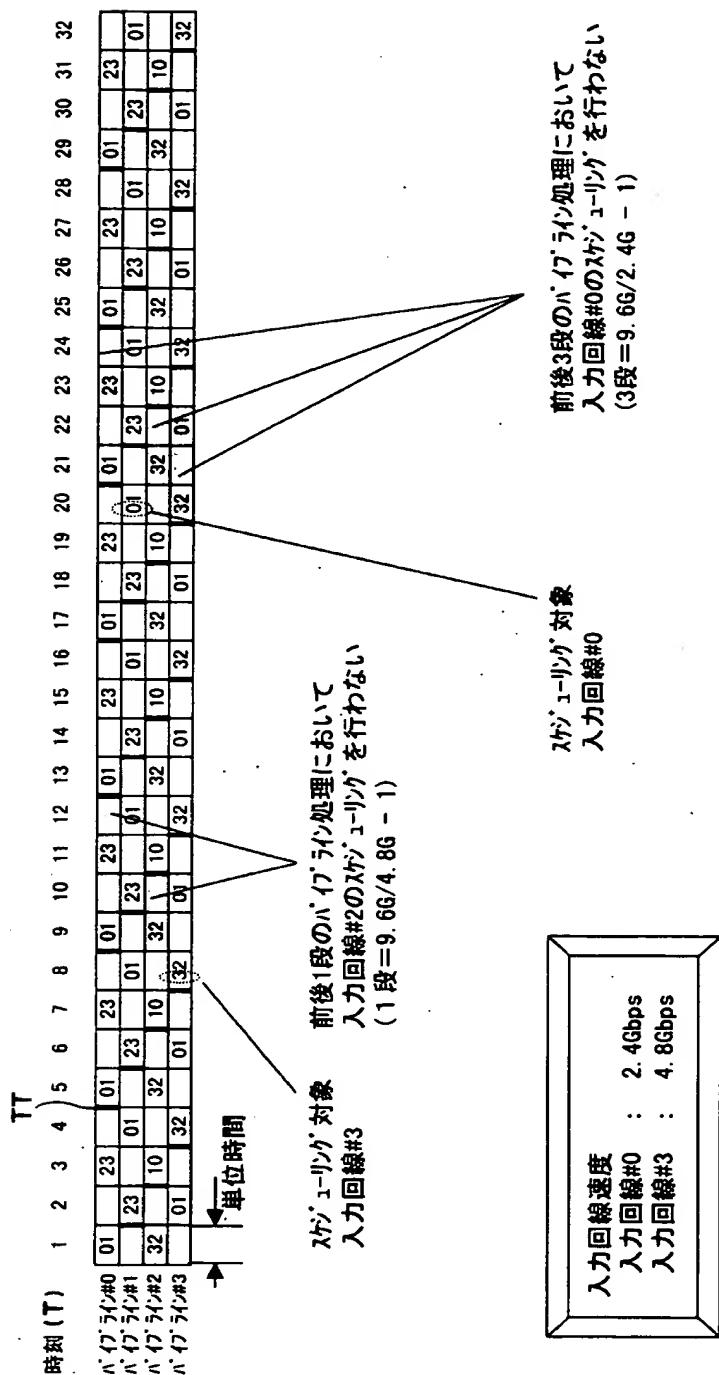
【図45】

第3のスケジューラの構成を示すブロック図



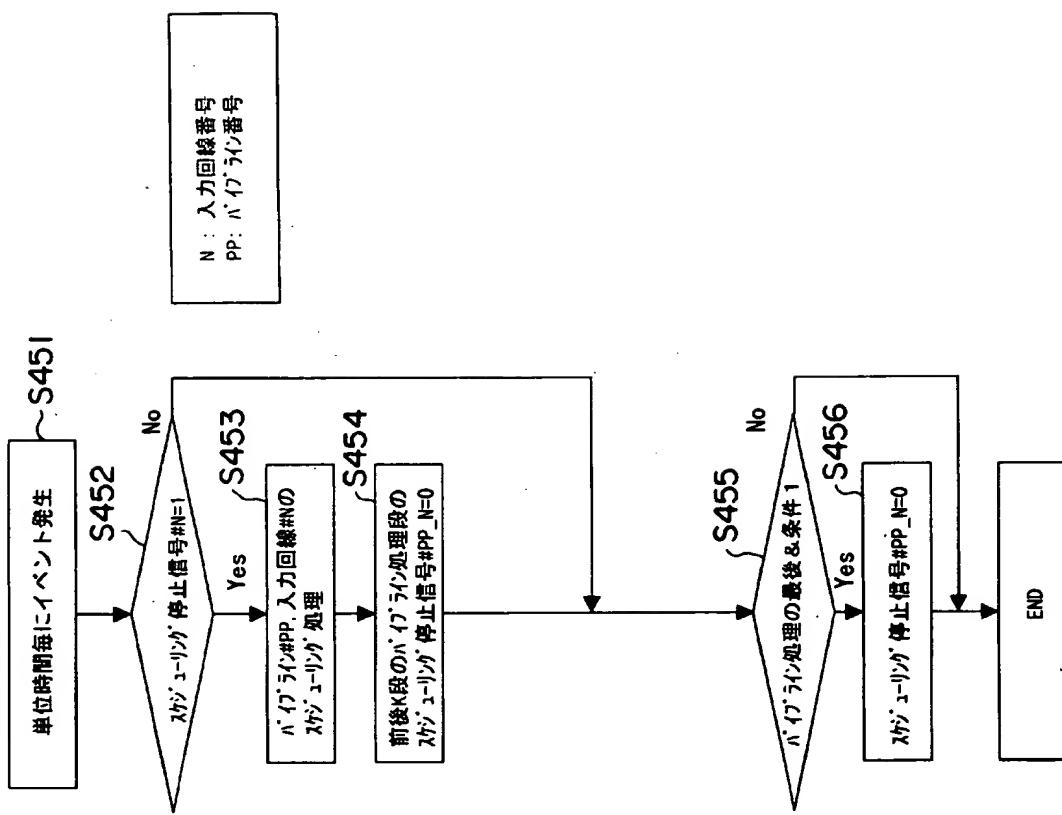
【図46】

入力回線速度に応じたスケジューリング処理を説明するための図



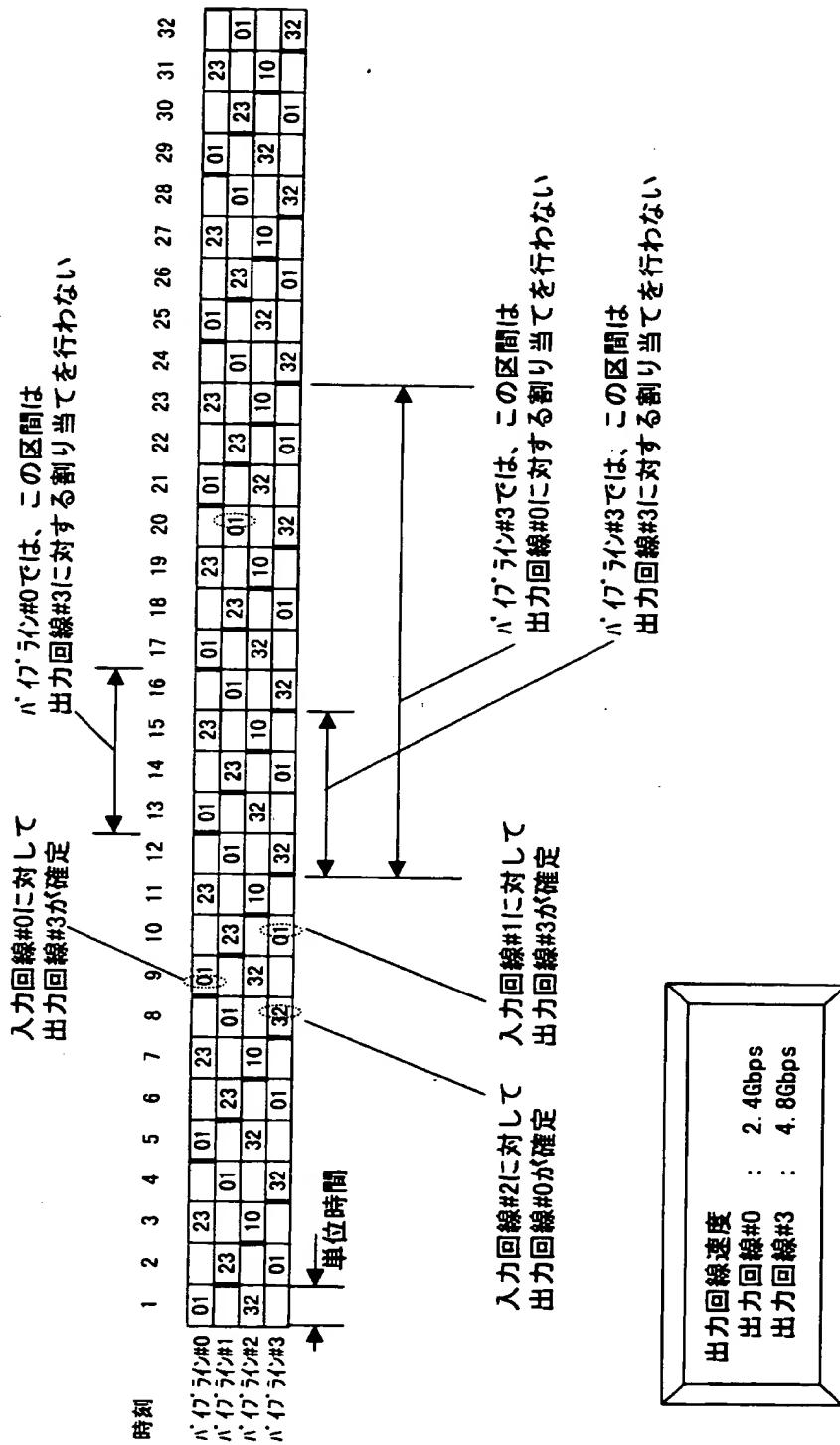
【図47】

入力回線速度に応じたスケジューリング処理手順を説明するための図

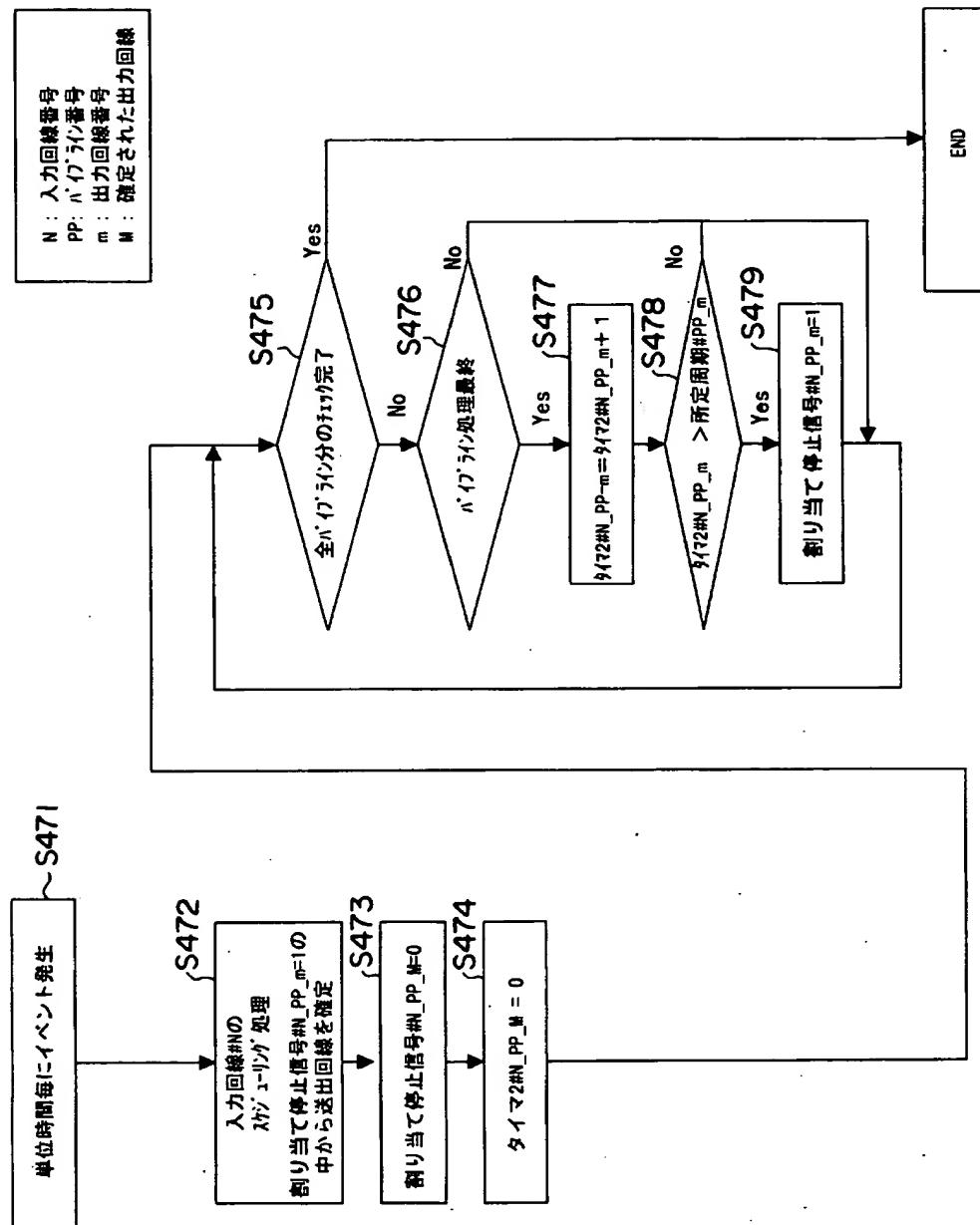


【図48】

出力回線速度に応じたスケジューリング処理を説明するための図

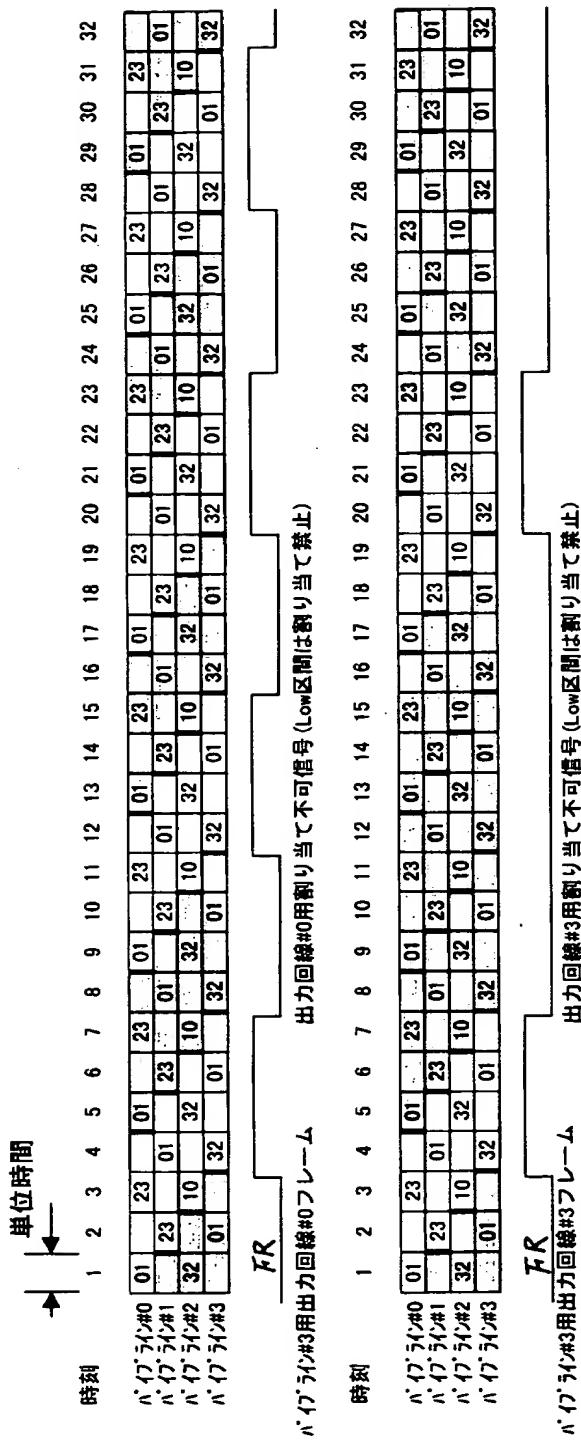


【図4.9】

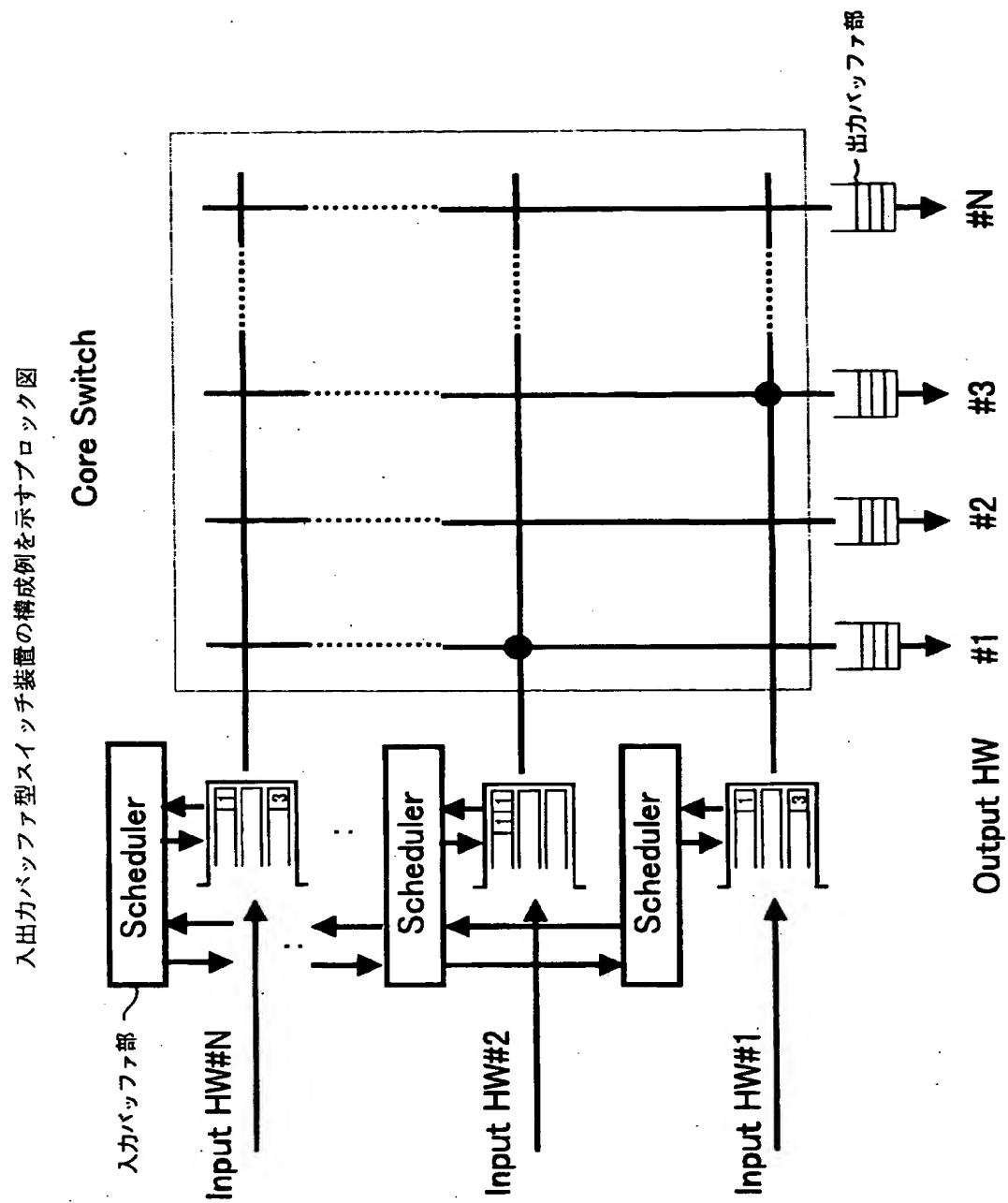


【図50】

第4のスケジューラにおける出力回線速度に応じたスケジューリング処理を説明するための図



【図51】



【書類名】 要約書

【要約】

【課題】 メモリアドレスをブロック単位で管理し、ブロック内の個別のアドレスは書き込み時または読み出し時にキュー毎に与えることでメモリ量を削減するなど。

【解決手段】 パケットスイッチ装置は、出力回線対応の論理的なキューを有する入力バッファメモリ手段と、スケジューリング開始入力回線を示す第1のポインタの制御手段と、スケジューリング対象回線のスケジューリング開始出力回線を示す第2のポインタの制御手段と、所望の出力回線に対する送出要求情報を保持する要求管理体制と、複数の送出要求情報の中から前記第2のポインタが示す出力回線から検索を開始し、他の入力回線に確保されていない出力回線を選択するスケジューリング処理手段と、複数の固定長パケットを一時的に保持し順次に出力するパケットバッファメモリ手段と、前記パケットバッファメモリ手段から出力された固定長パケットをスイッチングするスイッチ手段と、前記パケットバッファメモリ手段のアドレスを複数パケット分の固定長ブロックに分け、アドレス管理をブロック単位に行うアドレス管理手段とを備える。

【選択図】 図12

出願人履歴情報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号

氏 名 富士通株式会社